(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 April 2002 (18.04.2002)

PCT

(10) International Publication Number
WO 02/31111 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US01/27760

(22) International Filing Date: 11 October 2001 (11.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/687,527 12 October 2000 (12.10.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 09/687,527 (CIP)
Filed on 12 October 2000 (12.10.2000)

(71) Applicant *(for all designated States except US)*: **HYSEQ, INC.** [US/US]; 670 Almanor Avenue, Sunnyvalle, CA 94086 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **TANG, Y. Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **LIU, Chenghua** [CN/US]; 1125 Ranchero Way #14, San Jose, CA 95117 (US). **ZHOU, Ping** [US/US]; 7595 Newcastle Drive, Cupertino, CA 95014 (US). **ASUNDI, Vinod** [US/US]; 709 Foster City Boulevard, Foster City, CA 94404 (US). **ZHANG, Jie** [CN/US]; 4930 Poplar Terrace, Campbell, CA 95008 (US). **ZHAO, Qing, A.** [CN/US]; 1556 Kooser Road, San Jose, CA 95118 (US). **REN, Feiyan** [US/US]; 7703 Oak Meadow Court, Cupertino, CA 95014 (US). **XUE, Aidong, J.** [CN/US]; 1621 S. Mary Avenue, Sunnyvale, CA 94087 (US). **YANG, Yonghong** [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **WEHRMAN, Tom** [US/US]; CCSR Mol Pharm 3210, 269 W. Campus Drive, Stanford, CA 94305 (US). **DRMANAC, Radoje, T.** [US/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US).

(74) Agent: **ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris, Glovsky, and Popeo PC, One Financial Center, Boston, MA 02111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/31111 A2

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

2. BACKGROUND

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, circulating soluble factors, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize

one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-446. The polypeptides sequences are designated SEQ ID NO: 447-892. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is unknown or any of the four bases.

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO: 1-446 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO: 1-446. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO: 1-446 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO: 1-446. The sequence information can be a segment of any one of SEQ ID NO: 1-446 that uniquely identifies or represents the sequence information of SEQ ID NO: 1-446.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information are provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO: 1-446 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO: 1-446 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO: 1-446; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO: 1-446; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1-446. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO: 1-446; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in SEQ ID NO: 447-892; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a

nucleotide sequence set forth in SEQ ID NO: 1-446; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, *e.g., in situ* hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the

polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other

substances that interact with (*e.g.*, bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and detecting the complex by detecting the reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provide methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases or disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2); for which they have a signature region (as set forth in Table 3); or for which they have homology to a gene family (as set forth in Table 4). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

4. DETAILED DESCRIPTION OF THE INVENTION

4.1 DEFINITIONS

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the

natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived. The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or

synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 9 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NO: 1-446.

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular

Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO: 1-446. The sequence information can be a segment of any one of SEQ ID NO: 1-446 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO: 1-446. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4^{20} possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match ($1\div4^{25}$) times the increased probability for mismatch at each nucleotide position (3 x 25). The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements e.g. repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 500 amino acids, more preferably less than 200 amino acids more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include an initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

The term "variant"(or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, *e g.*, recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by

comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells

chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, *e.g.*, polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (*e.g.*, microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, *e.g., E. coli*, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include

an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (*e.g.*, soluble proteins) or partially (*e.g.*, receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2): 134 -143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (i.e., hybridization to filter-bound DNA in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (i.e., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (*i.e.*, the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, *e.g.*, mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, *e.g.*, mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% identity, more preferably at least 98% identity, and most preferably at least 99% identity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, and most preferably at least about 95% identity, more preferably at least about 98% sequence identity, and most preferably at least about 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence

(*e.g.*, via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J. (1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO: 1-446; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO: 447-892; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NO: 447-892. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO: 1-446; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing as SEQ ID NO: 447-892; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d)

a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 447-892. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO: 1-446 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO: 1-446 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO: 1-446 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, *e.g.*, at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least

about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99%, sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO: 1-446, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that are selective for (i.e. specifically hybridize to) any one of the polynucleotides of the invention are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided in SEQ ID NO: 1-446, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO: 1-446 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO: 1-446, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, *e.g.*, by substituting first with conservative choices (*e.g.*, hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (*e.g.*, hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., *DNA* 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith,

Nucleic Acids Res. 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and *Current Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO: 1-446, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et

al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO: 1-446 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO: 1-446 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.*, the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, *e.g.*, stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced

or derepressed by appropriate means (*e.g.*, temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

4.3 ANTISENSE NUCLEIC ACIDS

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1-446, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID NO: 447-892 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO: 1-446 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding a nucleic acid disclosed herein (*e.g.*, SEQ ID NO: 1-446), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of an mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or

genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be designed based upon the nucleotide sequence of a DNA disclosed herein (*i.e.*, SEQ ID NO: 1-446). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is

complementary to the nucleotide sequence to be cleaved in an mRNA of SEQ ID NO: 1-446 (see, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742). Alternatively, polynucleotides of the invention can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (*e.g.*, promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, Proc. *Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition,

Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5′ flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the

control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No.

PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO: 447-892 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO: 1-446 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO: 1-446 or (b) polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO: 447-892 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO: 447-892 or the corresponding full length or mature protein; and "substantial equivalents" thereof (*e.g.*, with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO: 447-892.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide

sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, *e.g.*, pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (*e.g.*, an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic

sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, *e.g.*, Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: *A Laboratory Manual*; Ausubel et al., *Current Protocols in Molecular Biology*. Polypeptide fragments that retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for e.g., small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO: 447-892.

The protein of the invention may also be expressed as a product of transgenic animals, *e.g.*, as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, *e.g.*, U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, Calif., U.S.A. (the MaxBat™ kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (*i.e.*, from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl™ or Cibacrom blue 3GA Sepharose™; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, *e.g.*, silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, e.g., targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, e.g., antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes,

dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference), the GeneAtlas software (Molecular Simulations Inc. (MSI), San Diego, CA) (Sanchez and Sali (1998) Proc. Natl. Acad. Sci., 95, 13597-13602; Kitson DH et al, (2000) "Remote homology detection using structural modeling – an evaluation" Submitted; Fischer and Eisenberg (1996) Protein Sci. 5, 947-955), Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark), and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a

fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.
In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprise one or more domains fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e,g.,* cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs

between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

4.8 GENE THERAPY

Mutations in the polynucleotides of the invention may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected *ex vivo, in situ*, or *in vivo* by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered *in vivo* to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.9 TRANSGENIC ANIMALS

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals,

can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the

polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

4.10 USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

4.10.1 RESEARCH USES AND UTILITIES

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant

protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

4.10.2 NUTRITIONAL USES

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent

stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells *in vivo* or *ex vivo* is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for

generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell sources (including hematopoietic stem cells and embryonic stem cells) and

cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support e.g. as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation,

those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

4.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of

bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as

stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and

disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (*e.g.*, anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastbom et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of

an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In

addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β_2 microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation,

those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

4.10.8 ACTIVIN/INHIBIN ACTIVITY

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present

invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of

cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostasis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the

invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without

necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine. Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These *in vitro* models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-

97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

4.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide

to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14 . Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

4.10.13 DRUG SCREENING

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science 282*:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis

methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, *Curr. Opin. Biotechnol.* 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., *Mol. Biotechnol,* 9(3):205-23 (1998); Hruby et al., *Curr Opin Chem Biol,* 1(1):114-19 (1997); Dorner et al., *Bioorg Med Chem,* 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (*i.e.*, increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population

expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications i.e. phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this

invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

4.10.16 LEUKEMIAS

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

(i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;

(ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

(iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;

(iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;

(v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;

(vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;

(vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and

(viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

(ii) increased sprouting of neurons in culture or in vivo;

(iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or

(iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set

forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, *etc*., depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, *e.g.*, weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of

the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

4.10.19 IDENTIFICATION OF POLYMORPHISMS

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

4.10.20 ARTHRITIS AND INFLAMMATION

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

4.11 THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

4.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An

exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01 μg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1 μg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth factor (PDGF), transforming

growth factors (TGF-α and TGF-β), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, *e.g.*, treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the

present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co- administered with one or more cytokines, lymphokines or other hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

4.12.1 ROUTES OF ADMINISTRATION

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated

from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or

other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene

glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, *e.g.*, gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable

polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with

inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not

increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 μg to about 100 mg (preferably about 0.1 μg to about 10 mg, more preferably about 0.1 μg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole

weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, *e.g.*, amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (*e.g.*, bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by

periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC_{50} as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD_{50} and ED_{50}. Compounds which exhibit high therapeutic

indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, *e.g.*, Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from *in vitro* data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 μg/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 μg/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

4.12.4 PACKAGING

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be

prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

4.13 ANTIBODIES

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, i.e., molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab}, $F_{ab'}$ and $F_{(ab')2}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG_1, IgG_2, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, such as the amino acid sequences shown in SEQ ID NO: 447-892, and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues, or at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, *e.g.*, a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for

targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, *Proc. Nat. Acad. Sci. USA* 78: 3824-3828; Kyte and Doolittle 1982, *J. Mol. Biol.* 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

4.13.1 POLYCLONAL ANTIBODIES

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents.

Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

4.13.2 MONOCLONAL ANTIBODIES

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly

myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal. The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

4.13.3 HUMANIZED ANTIBODIES

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the

imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

4.13.4 HUMAN ANTIBODIES

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein. Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild et al,(Nature Biotechnology 14, 845-51 (1996)); Neuberger (Nature

Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse™ as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in

culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

4.13.5 F_{ab} FRAGMENTS AND SINGLE CHAIN ANTIBODIES

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab')2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab')2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

4.13.6 BISPECIFIC ANTIBODIES

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the

correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. $F(ab')_2$ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate $F(ab')_2$ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB

derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody $F(ab')_2$ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., J. Immunol. 147:60 (1991).
Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific

antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

4.13.7 HETEROCONJUGATE ANTIBODIES

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

4.13.8 EFFECTOR FUNCTION ENGINEERING

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

4.13.9 IMMUNOCONJUGATES

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi, ^{131}I, ^{131}In, ^{90}Y, and ^{186}Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such as streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

4.14 COMPUTER READABLE SEQUENCES

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media"

refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (*e.g.* text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO: 1-446 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO: 1-446 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important

proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

4.15 TRIPLE HELIX FORMATION

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA. Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem: 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

4.16 DIAGNOSTIC ASSAYS AND KITS

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with

nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the

following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

4.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide *in vivo* at the target site.

4.18 SCREENING ASSAYS

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO: 1-446, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

(a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and

(b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein

encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the

ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

4.19 USE OF NUCLEIC ACIDS AS PROBES

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO: 1-446. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from any of the nucleotide sequences SEQ ID NO: 1-446 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent *in situ* hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

Oligonucleotides, i.e., small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata *et al.*, 1985; Dahlen *et al.*, 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller *et al.*, 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, e.g., Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound

to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen *et al.*, (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/μl) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm$_7$), is then added to a final concentration of 10 mM 1-MeIm$_7$. The single-stranded DNA solution is then dispensed into CovaLink NH strips (75 μl/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm$_7$, is made fresh and 25 μl added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may

also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The

results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, *CviJI*, described by Fitzgerald *et al.* (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease *CviJI* normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme (*CviJI***), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a *CviJI*** digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that *CviJI*** restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 μg instead of 2-5 μg); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

4.22 PREPARATION OF DNA ARRAYS

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm^2, depending on the type

of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane. Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm^2 and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

5. EXAMPLES

5.1 EXAMPLE 1

<u>Novel Nucleic Acid Sequences Obtained From Various Libraries</u>

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems (ABI) sequencer to obtain the novel nucleic acid sequences

5.2 EXAMPLE 2

Assemblage of Novel Nucleic Acids

The nucleic acids of the present invention, designated as SEQ ID NO: 1-446 were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST, gb pri, UniGene, and exons from public domain genomic sequences predicated by GenScan) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Further, inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), full-length gene sequences and their corresponding protein sequences were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTXY algorithm against Genbank (i.e., dbEST, gb pri, UniGene, and Genpept). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.). The full-length nucleotide sequences are shown in the Sequence Listing as SEQ ID NO: 1-446. The corresponding polypeptide sequences are SEQ ID NO: 447-892.

Table 1 shows the various tissue sources of SEQ ID NO: 1-446.

The nearest neighbor results for polypeptides encoded by SEQ ID NO: 1-446 (i.e. SEQ ID NO: 447-892) were obtained by a BLASTP (version 2.0al 19MP-WashU) search against Genpept, Geneseq and SwissProt databases using BLAST algorithm. The nearest neighbor result showed the closest homologue with functional annotation for SEQ ID NO: 1-446. The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The homologues with identifiable functions for SEQ ID NO: 1-446 are shown in Table 2 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), polypeptides encoded by SEQ ID NO: 1-446 (i.e. SEQ ID NO: 447-892) were examined to determine whether they had identifiable signature regions. Table 3 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the Pfam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) polypeptides encoded by SEQ ID NO: 1-446 (i.e. SEQ ID NO: 447-892) were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the p-value and the pFam score for the identified domain within the sequence.

The GeneAtlas™ software package (Molecular Simulations Inc. (MSI), San Diego, CA) was used to predict the three-dimensional structure models for the polypeptides encoded by SEQ ID NO: 1-446 (i.e. SEQ ID NO: 447-892). Models were generated by (1) PSI-BLAST which is a multiple alignment sequence profile-based searching developed by Altschul et al, (Nucl. Acids. Res. 25, 3389-3408 (1997)), (2) High Throughput Modeling (HTM) (Molecular Simulations Inc. (MSI) San Diego, CA,) which is an automated sequence and structure searching procedure (http://www.msi.com/), and (3) SeqFold™ which is a fold recognition method described by Fischer and Eisenberg (J. Mol. Biol. 209, 779-791 (1998)). This analysis was carried out, in part, by comparing the polypeptides of the invention with the known NMR (nuclear magnetic resonance) and x-ray crystal three-dimensional structures as templates. Table 5 shows, "PDB ID", the Protein DataBase (PDB) identifier given to template structure; "Chain ID", identifier of the subcomponent of the PDB template structure; "Compound Information", information of the PDB template structure and/or its subcomponents; "PDB Function Annotation" gives function of the PDB template as annotated by the PDB files (http:/www.rcsb.org/PDB/); start and end amino acid position of

the protein sequence aligned; PSI-BLAST score, the verify score, the SeqFold score, and the Potential(s) of Mean Force (PMF). The verify score is produced by GeneAtlas™ software (MSI), is based on Dr. Eisenberg's Profile-3D threading program developed in Dr. David Eisenberg's laboratory (US patent no. 5,436,850 and Luthy, Bowie, and Eisenberg, Nature, 356:83-85 (1992)) and a publication by R. Sanchez and A. Sali, Proc. Natl. Acad. Sci. USA, 95:13597-12502. The verify score produced by GeneAtlas normalizes the verify score for proteins with different lengths so that a unified cutoff can be used to select good models as follows:

Verify score (normalized) = (raw score – 1/2 high score)/(1/2 high score)

The PFM score, produced by GeneAtlas™ software (MSI), is a composite scoring function that depends in part on the compactness of the model, sequence identity in the alignment used to build the model, pairwise and surface mean force potentials (MFP). As given in Table 5, a verify score between 0 to 1.0, with 1 being the best, represents a good model. Similarly, a PMF score between 0 to 1.0, with 1 being the best, represents a good model. A SeqFold™ score of more than 50 is considered significant. A good model may also be determined by one of skill in the art based all the information in Table 5 taken in totality.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determined from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et al, as reference, were obtained for the polypeptide sequences. Table 6 shows the position of the last amino acid of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide.

Table 7 correlates each of SEQ ID NO: 1-446 to a specific chromosomal location.

Table 8 is a correlation table of the novel polynucleotide sequences SEQ ID NO: 1-446, and their corresponding priority nucleotide sequences in the priority application USSN 09/687,527, herein incorporated by reference in its entirety.

TABLE 1

Tissue Origin	RNA/Tissue Source	Library Name	SEQ ID NO:
adult brain	GIBCO	AB3001	31 35-36 52-53 57-59 63-64 69 73-74 86-87 102 109 138 140 148 151 153 163 177 179 194 235 240 250 276
adult brain	GIBCO	ABD003	2 6-8 10 13 20-22 35-36 38-39 43 45 51-54 58 60 64 68-69 73-74 76 86-87 90 93-94 97 100 109 117 120-123 127-128 131 137-138 140 145 148-149 151 155 159 163-164 166-167 170 172 174 179 181 187 189 196 199 207 209 211-212 232 238 245 259 262-263 267 269 276-277 305 324 337 341 418
adult brain	Clontech	ABR001	34 40 93 97 130 155 160 190 276 307 341 436
adult brain	Clontech	ABR006	15 30 61 65 68 70 74 88 95 99 106 109 113 129 134 138 148-149 151 154 160 179 190 200 207 210 219 228 231 240 248 250 267 275 284 315 317 335 355 373 401 415 426-428
adult brain	Clontech	ABR008	1 3-5 8-10 22 26 28-29 33-34 37 42 46 51 55-56 58 62-63 65 67-69 72 81 84-85 90 93 97-99 112-114 119 121-122 126-127 129 132 134-135 137 143-144 149-151 153-156 160 162 172 174 182-183 187 190-191 194-196 202 204-205 207 209-210 212 217 225-228 231 234 237-238 241-243 245 254 259-260 262 268 270 272-274 276 278-279 282 290 293-294 299 302 304 306 311 315-316 324 329 334 336 341-343 355 358-363 373-374 376-377 379 381-382 393 401-402 415 422 432 434-436
adult brain	Clontech	ABR011	52 155 160 315
adult brain	BioChain	ABR012	64 67 164 284
adult brain	BioChain	ABR013	356
adult brain	Invitrogen	ABR014	58 122 128 174 212 231 248 260
adult brain	Invitrogen	ABR015	6-7 58 63 72 80 122-123 269
adult brain	Invitrogen	ABR016	20-21 36 58 93 131 167 217 285
adult brain	Invitrogen	ABT004	13 33 36 58 63 75 93 95 99 102 107 120-121 123 127 143 149 154-155 160 166 179 185 189 202 208 210 212 219 222 228 235 237-238 250 259 269 272-274 276 279-280 282 294 306 312-313 317-318 324-325 329 402 436
cultured preadipocytes	Stratagene	ADP001	34-37 55 60 67 80 86-87 106 109 158 179-180 222 242 270 280 414
adrenal gland	Clontech	ADR002	8 19-21 25 36 42 44-45 47 55 59 62 68 72-73 84 87-88 114 121 127 144 149 152 179 181 202 204 217 225 248 263 292-293 321 357 415 433
adult heart	GIBCO	AHR001	6-9 15 19-21 30 33-36 39 43 45 49-51 53-55 57-59 61-64 67-70 73 75 80 84 86-87 95 97-98 100 103-104 109 112 114-115 117-118 125-126 128 131 134 136-137 139-140 145-146 149-152 158 162-163 174-175 177-179 181 184-186 193 196 200 202 205 207-210 213-220 228 241 243 245 248-249 255 263 269 276 278-279 287 289 291 296-297 299 302 305 308 330-332 382 393 402 425 432
adult kidney	GIBCO	AKD001	6-8 10 12 15 20-22 25-26 28-30 33-34 36-37 39 43-45 48 53-55 57-58 60 62-64 67-68 70-73 75-76 80-81 84 86-88 90 94-97 102-104 107 109 112 114-116 118 120-124 126-129 131 134 140 145 147 149 151-153 158 160-165 174-179 181-182 187-190 194-196 198 202 206 210-212 217-231 234-236 238 240 245-247 250-254 262-263 267 269-271 284 300 341 417 432
adult kidney	Invitrogen	AKT002	3-4 6-9 13 23 28-29 34 36 61-63 68 70 76 95 97-99 115 120-121 124 127-128 135 156-157 161 163 172 177 182 189 200 212 219 225 228 233 243-244 248 254-255 266

Tissue Origin	RNA/Tissue Source	Library Name	SEQ ID NO:
			271-274 281 303 316 321 323 334 347 400 417
adult lung	GIBCO	ALG001	6-8 34 40 53 58-59 64 67-68 73 76 109 112 118 129 134 136-137 153 159 163-164 175 179 187 191 193-194 196 200 208 235-236 240 243 251 255 263 275 317
lymph node	Clontech	ALN001	37 39 56 62-63 67 99 104 149 152 163-164 174 196 217 228 236 240 255 260 284 415
young liver	GIBCO	ALV001	20-21 33 54-55 59-61 72 76 88 95 100 115 121 123 125 127 137 141 149 158 170 172 179 186 194 196 200 209-210 221-222 226-227 240 244-245 251 256 258 263 269 432
adult liver	Invitrogen	ALV002	30 36 39 51-52 69 75 84 88 119-121 123 127 145 185 189 202 207 209-210 235 243 250 254 268-269 291 312 325 342 352 409 432
adult liver	Clontech	ALV003	26 80
adult ovary	Invitrogen	AOV001	2-4 6-10 12 15 19-23 25-26 28-30 32-34 36-39 42-43 47 51-54 56 59-65 67-68 71-73 75-76 78 84-88 90-94 97-98 102-104 108-110 113-114 116-117 119-121 123-128 131 136-142 145 149-150 152-153 155-156 159-164 172-176 178-181 183-184 187-189 191 193-196 200-202 207 209 212-213 217 219-220 222 228 231-238 240-241 243-247 250 253-255 257-259 262-263 265-267 269-270 272-274 280 283-284 294 302 306-311 319-320 322 330 333 335-336 341 350 409 415 417 431 436
adult placenta	Invitrogen	APL001	43 59 77 181 209
placenta	Invitrogen	APL002	10 22 24 34 36 73 77 121 285 300
adult spleen	GIBCO	ASP001	6-7 10 12 16-17 20-23 30 35 48 51 55 59 62-64 67 72 76 86-87 97 103-104 121 124 126 129 134 153 155 163-164 180-181 187 194 196 202 206 210 212 220 228 236 262 270 284 286-287 289 300 324 400 417
adult testis	GIBCO	ATS001	5 9 13 19-21 34 39 49-50 59 62 64 69 72 90 94-95 102 115 117-118 127 139 141 145 149 151-153 163 174-175 179 181 196 201 206-207 211 220 242 250 259 267 270
bladder	Invitrogen	BLD001	38 42 51 67 73 93 95 98 107 127 135 166 181 268 316
bone marrow	Clontech	BMD001	2 6-7 9 11-12 19-21 23 33-34 36 38 43 45 47 52 59 61-62 64 66 72-73 76 78 80 88 96 99-100 103-104 106 108 111-112 119 121 125 127-128 130 134-135 138 141 145 152-153 163-175 179 181 191 196 198-200 202-203 228 233-234 236 257 261 263 275 288 356 415 431-432 434-435 437-438
bone marrow	Clontech	BMD002	8-10 20-25 27 36-37 39-40 45 51-54 56-57 60-61 65-66 72 76 83-84 98 100 103-104 113 118-119 126 128 131-132 134 151 168-169 171-172 174 176 181 186 200 202-203 215 228 241-245 248 261 263 265 269-270 278-279 289 298-300 309 319 321 334-335 342 350 356 400 407 429 433-438 440
bone marrow	Clontech	BMD004	40 64 279
adult colon	Invitrogen	CLN001	27 48 58 100 122 128 157 179 185 212 246-247 317 355 384
mixture of 16 tissues/ mRNA*s	various vendors*	CTL016	103-104 323
mixture of 16 tissues/ mRNA*s	various vendors*	CTL021	64 179 260 323 445
adult cervix	BioChain	CVX001	3-4 6-7 9-10 12-13 20-23 25-26 30 36-37 39-40 43 45 47 51 53-54 56 58-59 61 63-64 66-67 71-72 75-76 78 84 90-

Tissue Origin	RNA/Tissue Source	Library Name	SEQ ID NO:
			92 94 97 100 103-104 110 114 118-120 123 128 131 136 138 140-141 149 153 157-158 163 170 174 179 181 184 186 196 198-199 202 208 212 225 231 238 240 257 267 270 285 288 295 301 305 311 335 338 340 356 364-365 383 394 415 425
diaphragm	BioChain	DIA002	215
endothelial cells	Stratagene	EDT001	2 5-13 19-23 28-30 32-37 39 42-43 45 52-53 55-60 62-65 68-69 73 76 80-81 84 86-88 91-92 94-96 98 103-104 106 109-110 114-115 119-122 124 126-128 131-132 134-137 139-141 153 161 163-165 167 170 172 175 177-180 182 185-187 190 193 196 198 202 206-207 211 216-219 222-224 232 237-238 240 243-244 246-247 252 255 258 262-263 270 272-274 284 289-290 292 299 315 318 341 380 415 417-418
esophagus	BioChain	ESO002	64 196 279
fetal brain	Clontech	FBR001	55 85 395
fetal brain	Clontech	FBR004	91-92 199-200 316
fetal brain	Clontech	FBR006	5 12 14 28-29 31 33-34 37 43 46 58 61-63 65 68 73-74 81 88 93 95 97 103-104 112 119-120 122-123 126-128 132 136 144 147 149 156 159 164 166 172 174-175 191 204 207 217 226-227 232 234 237-238 241-242 254 259-262 270 272-274 292-293 300 302 317 341-342 362 366-368 373-374 379 381 401-402 415 422 425-426 443-444
fetal brain	Clontech	FBRS03	112
fetal brain	Invitrogen	FBT002	5 10 22 31 33-34 42 52 55 58 64 66 73 75 84 98 107 109 120 122-123 127 133-134 136 138 140 147 155-156 160 166 180 185 190 196 209 238 254 260 270 294 313 317-318 324 326-329 334 341
fetal heart	Invitrogen	FHR001	64 67-69 86-87 90 202 206 213-215 217 225 245 272-274 285 292-293 336 434-435 437-438
fetal kidney	Clontech	FKD001	30 57 62 64 88 163 171 198 200 238 261 437-438
fetal kidney	Clontech	FKD002	146 156 176 255
fetal kidney	Invitrogen	FKD007	122 316
fetal lung	Clontech	FLG001	37 78 90 112 269 354
fetal lung	Invitrogen	FLG003	5 12 48 51 69 104 120 128 137 177 194 202 212 216 250 256 295 318 322 365 385
fetal lung	Clontech	FLG004	63 76 126
fetal liver-spleen	Columbia University	FLS001	1-15 18-50 52-58 60-113 115 118-120 122-124 126-128 131-132 134 136 142 144-145 149 153 158-159 162-165 168 172 176-187 189 191-194 196 200-206 209 216-217 219-220 222 226-227 232 235 245-247 255 259 261 272-274 284 289-293 296-298 300 309 323 337 351 361 363 375 394 400 406-407 410 415 419 431-432 436
fetal liver-spleen	Columbia University	FLS002	5 9-12 15 20-26 28-31 34-35 38-41 44 47 49-50 53-55 64 67-69 71-75 78-79 81 85-89 91-92 95 98-99 103-104 106 108-110 113 116-118 121 123-124 126 128 131 134 141-142 145 149 158 163-164 168 172 178-179 181-184 187 189 191-192 198-199 201-202 204 206 209 216-217 219-222 232 236 238 241 251 254 263 268 272-275 277 280 286 289 300 303 308 320 322 336-337 341 350-351 369 378-380 398 408-410 420-421 431-435
fetal liver-spleen	Columbia University	FLS003	1 12 36 61 74 78 88 111 125 174 221 291 378 433
fetal liver	Invitrogen	FLV001	10 13 22 31 33 36 41 60 69 84 114 120-121 126 164 219 221 238 269 312 315 323 418
fetal liver	Clontech	FLV002	261 313

Tissue Origin	RNA/Tissue Source	Library Name	SEQ ID NO:
fetal liver	Clontech	FLV004	16-17 36 53 68 80 86-87 132 171 179 183 204 221 272-274 292-293 336 369 400 409 432 437-438
fetal muscle	Invitrogen	FMS001	28-29 31 36 45 48 62 67 74 102 107 122 181 196 208 215 218 245 258 264 279-280 292 294 296-297 323-324 335 368 385-386 434-435
fetal muscle	Invitrogen	FMS002	5 23 38 51 61 85 90 102 108 151 174 183 187 189 204 210 212 219 260 278-280 292-293 309 341-342 359 362 373 436 441-442
fetal skin	Invitrogen	FSK001	8 11 23 30 36 45 48 51 53 58 60 64 67 70 73 81 84 86-87 90 95 100 102-104 106 114 116 118 121 127-128 132 134 143 148 157-159 168 172 174 178-179 181 183 185 189 194 205 207-208 235 238 241 243 246-247 250 258 264 268-271 280 285 288 294 299 308-309 316-317 338 352 354 387-389 395-396 402-405 434-435
fetal skin	Invitrogen	FSK002	8 31 39 67 79 86-87 90 97 118 168 174 181 203 207 216 219 222 226-227 229 248 251 269 299 319 341 373 388 396 415 422 432 434-435
fetal spleen	BioChain	FSP001	67 203 238
umbilical cord	BioChain	FUC001	15 20-21 33 36 38-39 51 54 59-60 63 67 71 73 76 79 90 97-98 103-104 109 117-118 120 128 134 137 140 149-152 159 164 172 181 189-190 192 194 196 213 225 228-229 238 241 263 266 280 282 289 305 323 331 344-345 368 372 406 427
fetal brain	GIBCO	HFB001	3-4 8-10 12 15 18-22 30-31 33-34 36 43 45 47 52 55 57-59 62-63 65 68-70 74 76 78 80 84 86-87 93-94 97-98 103-104 112 114-123 131-164 172 177-178 184-185 206 209 219 222 226-227 240 244-245 249 267 276 284 294-295 300 432
infant brain	Columbia University	IB2002	5 8-9 13 15-17 20-21 25-26 28-29 31 33 36 43 51-56 59 67-70 73 80 84 86-88 90 93 95 98 107 109-110 114 117-118 121 123-124 126-127 129 134-136 138 145 147-148 150-151 154-155 160 162 165-166 170 172 176 179 181-183 186-188 196 200 209 212 219 222 229 231-232 240 243 259-260 262-263 268-269 280 287 290 294 299 306 312-313 316 324 334 350 354 360 402 417 427 432
infant brain	Columbia University	IB2003	5 10-11 22 40 42 46 51-52 54 56 62 65 70 93 97-98 102 117 121 123 128 134-135 140 151 154 160 165 183 208 219 243 259 269 294 299 306 316-317 324 341 354 436
infant brain	Columbia University	IBM002	93 95 123 140 181
infant brain	Columbia University	IBS001	54 73 93 123 176 188 220 255-256 331
lung, fibroblast	Stratagene	LFB001	6-7 32 35 55 60 64 71 103-104 109-110 118 123 128 137 140 145 161 163 175 187 193 217 225 236 243 264 337 377 416
lung tumor	Invitrogen	LGT002	3-4 6-7 10-12 14-15 20-22 27 34 36 38-39 42 48 51-52 54-56 58-60 63 66 68-69 71 73 76 78 80 84 86-89 95 98 103-104 109 114 116-118 120 123-124 127-128 131 135 137 141 145 153 157 163 172 178-179 182 186-187 191-194 196 199 201 206 210 218-224 226-228 233 235-236 243 251 253 255 261 265 270-271 280 289-290 296-297 300 303 310 312 324 332 334-335 351 353 365 376 417 427 431
Lymphocytes	ATCC	LPC001	6-7 9 16-17 25 28-29 33 36 53 55 57 64 66 78 84 86-87 94-95 97 104 114 125 139 149 153 170 172 174 177 186 191 195 200 219 228 232-233 243 254 256 292-293 302

Tissue Origin	RNA/Tissue Source	Library Name	SEQ ID NO:
			310 342 345 378 398 400 411-413
Leukocyte	GIBCO	LUC001	6-8 12 16-17 19-21 23 25 28-30 33-34 36 38 40 42-43 45 49-51 55-56 58-66 68 71 75-76 78 80 84 86-88 94-95 97-100 102-104 111-116 119-120 124-125 128-129 131 138-139 141 145 147 149 152-153 158-159 161 163-164 172 175-179 181-182 184-185 187 189 193-197 200 203 206-207 209-211 216-217 219-220 222 233-245 250 255 262-263 265-267 270 275 284 286 298 300 307 351 361 397 415 431 436
Leukocyte	Clontech	LUC003	51 62 68 70 73 80 95 97 117 163 181 206 228 267 310 415
melanoma from cell line ATCC #CRL 1424	Clontech	MEL004	9 15 20-21 34 51-52 61 64 68 71 76 80-81 106 119 122 124 163 172 186-187 196 223-224 226-227 258 262 291 302 341 396 415
mammary gland	Invitrogen	MMG001	8 10 13 15 22 28-29 33-34 36-37 42 51-52 55 58 60 62-63 72-73 84-85 88 90 95 98 102 114 118-122 127 132 134-135 137-138 140 143 145 149 151-152 165-166 168 175 178-180 184-185 189 196 202 209-210 212 217 219 222 235 238 244 246-247 250-251 257 268-269 271-274 290 295 299-304 319-320 324 330 334 337-339 341-342 352 369 371 415
induced neuron cells	Stratagene	NTD001	9 36 45 68 73 76 97 106 112 119 126 132 137-139 160 179 264 306 341 376 401
retinoic acid-induced neuronal cells	Stratagene	NTR001	36 118 134 221 261 401 418
neuronal cells	Stratagene	NTU001	33 36 46 68 72 81 91-92 98 102 112 160 182 190-191 198 222 258 261 271 314 316 342 418 423
pituitary gland	Clontech	PIT004	20-21 36 55 65 68 137-138 148 162-163 170 196 341 356 430
placenta	Clontech	PLA003	12 30 67 194 302 417 436
prostate	Clontech	PRT001	9-10 22-23 29 36 38 43 112 118 128 136-137 140 151 163 177 185 189 209 233 250 255 268 282 335 346 354 415 434-435
rectum	Invitrogen	REC001	27 42 60 69-70 98 103-104 123 149 165 172 235 251 302 318 324 372 379 390 432
salivary gland	Clontech	SAL001	6-7 9 33 48 53 62 157 164 170 177 190 194 257 268 287 312 322 365 391-392
skin fibroblast	ATCC	SFB001	63 112
skin fibroblast	ATCC	SFB003	112
small intestine	Clontech	SIN001	9-10 12 22 30 33 36 40 45 52 55 72 78 84 90 95 114 117 119 123-124 127 129 134 136 149-151 163 176 181-182 193 196 206 232 236 251 287 318 324 334 350 432 439
skeletal muscle	Clontech	SKM001	3-4 6-7 20-21 64 103-104 120 153 176-177 179 187 191 215 278-279 330 386
skeletal muscle	Clontech	SKMS04	42
spinal cord	Clontech	SPC001	9 12 33-34 36 38-39 45 53 56 58 61 64 66 78 86-87 90 98 126 151 157-158 160 178-179 181 185 196 206 210 217 245 250 262 267 270 276 282 298 347 355 370 415 424
adult spleen	Clontech	SPLc01	25 125 136 138 168 171 176 275 416
stomach	Clontech	STO001	69 73 94 97 100 141 177 231 233 237 245 339 372 402 415
thalamus	Clontech	THA002	58 72 78 93 127 133 138 160 184 190 259 269 282 415

Tissue Origin	RNA/Tissue Source	Library Name	SEQ ID NO:
thymus	Clontech	THM001	6-7 9 12 16-17 19 33 42 59 61 64 76 78 91-92 104 139 153 158 161 163 168 172 174-175 177 179 189 198 202 222 231 237 239 243 272-274 299 321 332 356 394
thymus	Clontech	THMc02	6-7 9 12 14 16-17 19 28-29 37-38 47 51 53-54 62-63 73 83 88 91-92 109 113 126 133 151 156 158 163 171 176 179 181 185 190 194 198 200 206 219 226-228 231-232 234 239 242-243 259 261 265 272-274 290 309 356 373-374 397-399 434-435 437-438
thyroid gland	Clontech	THR001	3-4 6-7 9-10 12 20-22 25-26 30 36 39-40 42 47 53-54 59-60 62 64 68-69 71 76 85 88 94-95 98 104 106 108-109 113 116 118-121 124-126 131-132 137 153 158-159 163-164 168 170 174 180 189-191 194 196 199 202 207 209 211 217 221-222 232 236-238 240 244 248 250 254-255 257 259 269-271 280 298 302-303 310 320 326 337-338 347 356 371 377 415 417-419 436
trachea	Clontech	TRC001	6-7 36 59 78 127 152 190 240 251 257 270 272-274 281 299 301 348-349 351 365
uterus	Clontech	UTR001	3-4 59 118 123 137 177 217 219 244 270 306 311 316 340 357 372 431

* The 16 tissue/mRNAs and their vendor sources are as follows: 1) Normal adult brain mRNA (Invitrogen), 2) Normal adult kidney mRNA (Invitrogen), 3) Normal fetal brain mRNA (Invitrogen), 4) Normal adult liver mRNA (Invitrogen), 5) Normal fetal kidney mRNA (Invitrogen), 6) Normal fetal liver mRNA (Invitrogen), 7) normal fetal skin mRNA (Invitrogen), 8) human adrenal gland mRNA (Clontech), 9) Human bone marrow mRNA (Clontech), 10) Human leukemia lymphoblastic mRNA (Clontech), 11) Human thymus mRNA (Clontech), 12) human lymph node mRNA (Clontech), 13) human so\spinal cord mRNA (Clontech), 14) human thyroid mRNA (Clontech), 15) human esophagus mRNA (BioChain), 16) human conceptional umbilical cord mRNA (BioChain).

TABLE 2

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
447	AAB87354	Homo sapiens	22-MAY-2001 31-AUG-2000 Human gene 13 encoded secreted protein HFVJP07, SEQ ID NO:95.	731	100
448	X97490	Mus musculus	PNG protein	739	96
449	AK001950	Homo sapiens	FLJ11088 fis, clone PLACE1005287, weakly similar to INNER CENTROMERE PROTEIN.	1157	100
450	AK001950	Homo sapiens	FLJ11088 fis, clone PLACE1005287, weakly similar to INNER CENTROMERE PROTEIN.	790	73
451	AB044385	Homo sapiens	mRNA for transmembrane molecule with thrombospondin module, complete cds.	4492	99
452	BC005361	Homo sapiens	proteasome (prosome, macropain) subunit, alpha type, 4, clone MGC:12467 IMAGE:3685931, mRNA, complete cds.	1334	100
453	BC005361	Homo sapiens	proteasome (prosome, macropain) subunit, alpha type, 4, clone MGC:12467 IMAGE:3685931, mRNA, complete cds.	1098	100
454	AK001930	Homo sapiens	FLJ11068 fis, clone PLACE1004918, weakly similar to L-LACTATE DEHYDROGENASE M CHAIN (EC 1.1.1.27).	1742	100
455	AF151042	Homo sapiens	HSPC208	740	100
456	AL365512	Homo sapiens	human gene mapping to chomosome 22.	2511	99
457	AF279307	Homo sapiens	function 1B (ASF1B) mRNA, complete cds.	1075	99
458	AF212243	Homo sapiens	mRNA, complete cds.	1104	100
459	AAY13360	Homo sapiens	25-JUN-1999 16-SEP-1998 Amino acid sequence of protein PRO269.	2350	100
460	AAB65692	Homo sapiens	27-MAR-2001 26-MAY-2000 Novel protein kinase, SEQ ID NO: 220.	2758	96
461	AK001061	Homo sapiens	FLJ10199 fis, clone HEMBA1004850.	1305	100
462	AF042380	Homo sapiens	adaptor protein (Grf40) mRNA, complete cds.	1785	100
463	AF042380	Homo sapiens	adaptor protein (Grf40) mRNA, complete cds.	809	100
464	AL137422	Homo sapiens	cDNA DKFZp761A1623 (from clone DKFZp761A1623); partial cds.	410	98
465	AF220193	Homo sapiens	hypothalamus protein HT007 mRNA, complete cds.	1039	100
466	AAB60505	Homo sapiens	24-APR-2001 21-JUL-2000 Human cell cycle and proliferation protein CCYPR-53, SEQ ID NO:53.	3419	100
467	AAB69556	Homo sapiens	27-APR-2001 10-MAR-2000 Human Repro-EN-1.0 protein.	3315	99
468	AL365512	Homo sapiens	human gene mapping to chomosome 22.	2294	99
469	AAB90821	Homo sapiens	15-JUN-2001 02-OCT-2000 Human shear stress-response protein SEQ ID NO: 150.	3011	99
470	AF195821	Homo sapiens	(TNG2) mRNA, complete cds.	562	100
471	AK000399	Homo sapiens	FLJ20392 fis, clone KAIA4653.	2281	99
472	AK001240	Homo sapiens	FLJ10378 fis, clone NT2RM2002004, weakly similar to LA PROTEIN HOMOLOG.	1654	100

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
473	Y11339	Homo sapiens	for GalNAc alpha-2, 6-sialyltransferase I, long form.	3182	100
474	AAY90296	Homo sapiens	24-OCT-2000 11-JAN-2000 Human peptidase, HPEP-13 protein sequence.	936	100
475	AAY90296	Homo sapiens	24-OCT-2000 11-JAN-2000 Human peptidase, HPEP-13 protein sequence.	2087	99
476	AJ250193	Mus musculus	muscle protein 637	730	72
477	AK001706	Homo sapiens	FLJ10844 fis, clone NT2RP4001353.	959	100
478	AAB56847	Homo sapiens	13-MAR-2001 08-MAR-2000 Human prostate cancer antigen protein sequence SEQ ID NO:1425.	749	100
479	AB049952	Homo sapiens	mRNA for mitochondrial ribosomal protein S18a, complete cds.	1074	100
480	AK011757	Mus musculus	putative	589	100
481	BC012167	Homo sapiens	Similar to RIKEN cDNA 3110030K20 gene, clone MGC:20409 IMAGE:4637888, mRNA, complete cds.	899	99
482	AF038129	Ovis aries	polyubiquitin	771	100
483	AK012782	Mus musculus	putative	2562	92
484	AK001214	Homo sapiens	FLJ10352 fis, clone NT2RM2001152.	2770	100
485	AK021681	Homo sapiens	FLJ11619 fis, clone HEMBA1004131, moderately similar to SEPTIN 2 HOMOLOG.	2337	100
486	AJ252060	Homo sapiens	for TRABID protein (TRABID gene).	3796	100
487	AL137301	Homo sapiens	cDNA DKFZp434N1429 (from clone DKFZp434N1429); partial cds.	261	60
488	BC007588	Homo sapiens	Similar to RIKEN cDNA 2310012N15 gene, clone IMAGE:3342825, mRNA, partial cds.	1328	92
489	AB015335	Homo sapiens	mRNA, partial cds.	617	100
490	AAY66765	Homo sapiens	05-APR-2000 02-JUN-1999 Membrane-bound protein PRO1384.	1251	99
491	AC005154	Homo sapiens	clone RP4-777O23 from 7p14-p15, complete sequence.	994	100
492	AF155140	Homo sapiens	testicular RNA helicase mRNA, complete cds.	1902	99
493	AK001760	Homo sapiens	FLJ10898 fis, clone NT2RP5003492.	2575	99
494	BC007396	Homo sapiens	clone IMAGE:3834655, mRNA, partial cds.	1428	100
495	AK001374	Homo sapiens	FLJ10512 fis, clone NT2RP2000658.	2604	100
496	AK001374	Homo sapiens	FLJ10512 fis, clone NT2RP2000658.	1902	97
497	AK000507	Homo sapiens	FLJ20500 fis, clone KAT09159.	1189	100
498	AK000650	Homo sapiens	FLJ20643 fis, clone KAT02633.	1490	99
499	AK001766	Homo sapiens	FLJ10904 fis, clone OVARC1000013, weakly similar to APOPTOTIC PROTEASE ACTIVATING FACTOR 1.	2403	100
500	AF233224	Homo sapiens	protein FBS (FBS) mRNA, complete cds.	1698	100
501	BC005030	Homo sapiens	clone MGC:12628 IMAGE:3690254, mRNA, complete cds.	1853	100
502	AK001449	Homo sapiens	FLJ10587 fis, clone NT2RP2004042.	3440	100
503	AF326206	Homo sapiens	transcription factor mRNA, complete cds.	2149	99
504	AF220191	Homo sapiens	hypothalamus protein HSMNP1 mRNA, complete cds.	1099	100
505	AF155096	Homo sapiens	antigen mRNA, partial cds.	2008	98
506	BC008250	Homo sapiens	Similar to RIKEN cDNA 0610025L15	1332	100

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			gene, clone MGC:9282 IMAGE:3872059, mRNA, complete cds.		
507	AAY90287	Homo sapiens	24-OCT-2000 11-JAN-2000 Human peptidase, HPEP-4 protein sequence.	2514	100
508	BC000802	Homo sapiens	Similar to ribosomal protein S9, clone MGC:5482 IMAGE:3452221, mRNA, complete cds.	976	99
509	BC010011	Homo sapiens	Similar to RIKEN cDNA 4931428D14 gene, clone MGC:15407 IMAGE:4309613, mRNA, complete cds.	1691	97
510	AK010605	Mus musculus	putative	1023	99
511	AK020026	Mus musculus	putative	799	97
512	AB051853	Homo sapiens	gene for rho-GTPase activating protein, complete cds.	1119	100
513	J04695	Mus musculus	alpha-2 type IV collagen	4444	87
514	AL512733	Homo sapiens	cDNA DKFZp762P093 (from clone DKFZp762P093).	1380	100
515	AF284752	Homo sapiens	(GK004) mRNA, complete cds.	654	100
516	AK001055	Homo sapiens	FLJ10193 fis, clone HEMBA1004763.	767	100
518	AB037669	Homo sapiens	mRNA for L-type amino acid transporter 2, complete cds.	2790	100
519	AC007059	Homo sapiens	19, cosmid R26549, complete sequence.	4163	100
520	AF119863	Homo sapiens	PRO2160	483	100
521	AF151063	Homo sapiens	HSPC229	984	100
522	BC012123	Homo sapiens	golgi phosphoprotein 3, clone MGC:20187 IMAGE:4558305, mRNA, complete cds.	1528	100
523	BC002717	Homo sapiens	Similar to chorionic somatomammotropin hormone 1 (placental lactogen), clone MGC:3714 IMAGE:3631916, mRNA, complete cds.	1128	100
524	AAB36522	Homo sapiens	07-MAR-2001 13-APR-2000 Human CLASP related protein sequence SEQ ID NO:4.	3431	99
525	AAB08763	Homo sapiens	02-JAN-2001 29-FEB-2000 A human leukocyte and blood related protein (LBAP).	608	99
526	BC001810	Homo sapiens	enolase 1, (alpha), clone MGC:2414 IMAGE:2906988, mRNA, complete cds.	2206	99
527	AAY87267	Homo sapiens	11-MAY-2000 25-JUN-1999 Human signal peptide containing protein HSPP-44 SEQ ID NO:44.	1790	100
528	AK002043	Homo sapiens	FLJ11181 fis, clone PLACE1007460.	682	100
529	AK001795	Homo sapiens	FLJ10933 fis, clone OVARC1000605.	891	100
530	AF182412	Homo sapiens	(MDS025) mRNA, complete cds.	1104	98
531	AF345564	Homo sapiens	FKSG76	1327	99
532	AK008759	Mus musculus	putative	1314	96
533	BC010543	Homo sapiens	clone MGC:17544 IMAGE:3462146, mRNA, complete cds.	1093	100
534	AK023510	Homo sapiens	FLJ13448 fis, clone PLACE1002993.	1257	100
535	AL110245	Homo sapiens	cDNA DKFZp434F011 (from clone DKFZp434F011); partial cds.	301	91
536	U04520	Homo sapiens	IV collagen alpha 5 chain (COL4A5) gene, exon 51 and complete cds, alternatively spliced.	3630	100
537	AK000516	Homo sapiens	FLJ20509 fis, clone KAT09623.	1088	100

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
538	AK000516	Homo sapiens	FLJ20509 fis, clone KAT09623.	788	100
539	AC004865	Homo sapiens	clone RP4-728D4, complete sequence.	3759	100
540	AF286162	Homo sapiens	4-phosphate Adaptor Protein-1 mRNA, complete cds.	1570	99
541	AL591714	Homo sapiens	human gene mapping to chomosome 20.	821	100
542	AX179297	Homo sapiens	21615 ADH	1243	100
543	AL136844	Homo sapiens	cDNA DKFZp434G1730 (from clone DKFZp434G1730); complete cds.	1583	100
544	AK001371	Homo sapiens	FLJ10509 fis, clone NT2RP2000617.	3677	100
545	AK000213	Homo sapiens	FLJ20206 fis, clone COLF1582.	2343	100
546	AK008020	Mus musculus	putative	1919	71
547	AF119870	Homo sapiens	PRO2266	616	100
548	AK000763	Homo sapiens	FLJ20756 fis, clone HEP01556.	3152	100
549	AK023550	Homo sapiens	FLJ13488 fis, clone PLACE1003915, weakly similar to PROBABLE ARGINYL-TRNA SYNTHETASE, CYTOPLASMIC (EC 6.1.1.19).	1215	99
550	AK023550	Homo sapiens	FLJ13488 fis, clone PLACE1003915, weakly similar to PROBABLE ARGINYL-TRNA SYNTHETASE, CYTOPLASMIC (EC 6.1.1.19).	2069	99
551	AC002126	Homo sapiens	from chromosome 19-cosmids R30102:R29350:R27740 containing MEF2B, genomic sequence, complete sequence.	449	100
552	BC012182	Homo sapiens	clone MGC:20469 IMAGE:4554554, mRNA, complete cds.	1582	99
553	AL136528	Homo sapiens	DNA sequence from clone RP5-1092A11 on chromosome 1p36.2-36.33 Contains the gene for KIAA0495 protein, the TP73 (tumor protein p73) gene, a gene containing a WD repeat domain, ESTs, STSs, GSSs and CpG Islands, complete sequence.	271	100
554	AF334161	Homo sapiens	finger protein mRNA, complete cds.	1561	98
555	AJ277587	Homo sapiens	mRNA for Spir-1 protein (Spir-1 gene).	3012	99
556	AY014283	Homo sapiens	mRNA, complete cds.	1066	100
557	AF090938	Homo sapiens	HQ0628 PRO0628 mRNA, complete cds.	278	100
558	AF161511	Homo sapiens	HSPC162	480	100
559	AF039942	Homo sapiens	transcription factor Zhangfei (ZF) mRNA, complete cds.	1382	100
560	AF271782	Homo sapiens	mRNA, complete cds.	1280	100
561	AF107495	Homo sapiens	and putative FWP002 mRNA, complete cds.	783	100
562	AK015086	Mus musculus	putative	183	70
563	AL353936	Homo sapiens	cDNA DKFZp761K1423 (from clone DKFZp761K1423).	533	100
564	X87241	Homo sapiens	mRNA for hFat protein.	19971	99
565	BC004896	Homo sapiens	Similar to ribosomal protein, mitochondrial, L5, clone MGC:3400 IMAGE:3529006, mRNA, complete cds.	1494	100
566	AB062594	Bos taurus	putative	704	87
567	AL136683	Homo sapiens	cDNA DKFZp564D0478 (from clone DKFZp564D0478); complete cds.	1034	100
568	AAY87355	Homo sapiens	11-MAY-2000 25-JUN-1999 Human	952	100

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			signal peptide containing protein HSPP-132 SEQ ID NO:132.		
569	BC008967	Homo sapiens	clone IMAGE:3010666, mRNA, partial cds.	1024	100
570	AK022754	Homo sapiens	FLJ12692 fis, clone NT2RM4002623, weakly similar to ASPARTYL-TRNA SYNTHETASE (EC 6.1.1.12).	2425	99
571	AF083106	Homo sapiens	type 1 (SIRT1) mRNA, complete cds.	3929	100
572	AK000017	Homo sapiens	FLJ20010 fis, clone ADKA03268.	611	100
573	AF308801	Homo sapiens	protein sorting protein 16 (VPS16) mRNA, complete cds.	2541	99
574	BC001686	Homo sapiens	methionine adenosyltransferase II, alpha, clone MGC:2907 IMAGE:3010820, mRNA, complete cds.	1315	98
575	AK000675	Homo sapiens	FLJ20668 fis, clone KAIA585.	1474	100
576	X68242	Homo sapiens	mRNA for Hin-1.	757	100
577	BC001245	Homo sapiens	Similar to uncharacterized bone marrow protein BM036, clone MGC:4957 IMAGE:3460193, mRNA, complete cds.	1504	99
578	BC009782	Homo sapiens	hypothetical protein dJ122O8.2, clone MGC:13493 IMAGE:4092710, mRNA, complete cds.	432	98
579	AL133109	Homo sapiens	cDNA DKFZp566N1047 (from clone DKFZp566N1047); partial cds.	3416	99
580	AF161494	Homo sapiens	HSPC145	1562	100
581	AAY22465	Homo sapiens	29-SEP-1999 17-DEC-1998 Human hippocampal sel-10 protein sequence.	216	23
582	AF312864	Homo sapiens	mRNA, complete cds.	627	100
583	AAY70236	Homo sapiens	06-JUN-2000 20-AUG-1999 Human RNA-associated protein-17 (RNAAP-17).	2310	100
584	AF240769	Macaca mulatta	VAMP-2	584	100
585	AAB98084	Homo sapiens	16-AUG-2001 26-OCT-2000 Human protein sequence SEQ ID NO:110.	2482	99
586	AK002058	Homo sapiens	FLJ11196 fis, clone PLACE1007688, weakly similar to LA PROTEIN HOMOLOG.	2551	99
587	AK000500	Homo sapiens	FLJ20493 fis, clone KAT08512.	834	100
588	AF251296	Homo sapiens	mRNA, complete cds.	1299	100
589	AF169149	Homo sapiens	(CABP1) mRNA, complete cds.	1172	99
590	M96859	Homo sapiens	dipeptidyl aminopeptidase like protein mRNA, complete cds.	2246	52
591	AAB88489	Homo sapiens	23-MAY-2001 07-JUL-2000 Human membrane or secretory protein clone PSEC0265.	967	100
592	AB063495	Mus musculus	Spred-1	2205	92
593	AF155661	Homo sapiens	dehydrogenase (PDH) mRNA, complete cds.	3050	100
594	AAY90962	Homo sapiens	05-SEP-2000 12-OCT-1999 Human G713 protein sequence SEQ ID NO:5.	1403	99
595	AF315378	Rattus norvegicus	suppressor of profilin/p41 of actin-related complex 2/3	1975	98
596	AB036693	Homo sapiens	for RAB9-like protein, complete cds.	1067	100
597	AF359284	Homo sapiens	mRNA, complete cds.	5004	99
598	AK001877	Homo sapiens	FLJ11015 fis, clone PLACE1003302,	2746	99

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			highly similar to ZINC FINGER PROTEIN 83.		
599	AAB71913	Homo sapiens	09-MAY-2001 16-AUG-2000 Human ISOM-5.	1516	100
600	L27867	Rattus norvegicus	neurexophilin	1448	98
601	AC004991	Homo sapiens	clone RP5-1186C1 from 7q21.2-q31.1, complete sequence.	311	100
602	AF057019	Dictyostelium discoideum	interaptin	146	26
603	AF247177	Mus musculus	sphingosine-1-phosphate phosphohydrolase	523	36
604	BC007704	Homo sapiens	clone MGC:10277 IMAGE:3952366, mRNA, complete cds.	746	100
605	U18920	Homo sapiens	chromosome 17q12-21 mRNA, clone pOV-3, partial cds.	455	81
606	U48363.	Mus musculus	alpha-NAC, muscle-specific form gp220	810	30
607	AF014008	Bos taurus	myocardial vascular inhibition factor	490	100
608	AL136604	Homo sapiens	cDNA DKFZp564F2122 (from clone DKFZp564F2122); complete cds.	2716	96
609	AK007689	Mus musculus	putative	289	100
610	AAB57020	Homo sapiens	13-MAR-2001 08-MAR-2000 Human prostate cancer antigen protein sequence SEQ ID NO:1598.	384	100
611	AAB20328	Homo sapiens	29-MAY-2001 14-SEP-2000 Human protein phosphatase and kinase protein-7.	798	100
612	BC007618	Homo sapiens	clone MGC:15730 IMAGE:3355289, mRNA, complete cds.	2163	100
613	AAZ94941	Homo sapiens	01-AUG-2000 29-SEP-1999 Human carbohydrate-associated protein CRBAP-1 cDNA.	654	100
614	M91669	Homo sapiens	Bullous pemphigoid autoantigen BP180 gene, 3' end.	8016	99
615	AF116649	Homo sapiens	PRO0566	248	100
616	AK001837	Homo sapiens	FLJ10975 fis, clone PLACE1001383, weakly similar to ZINC-FINGER PROTEIN UBI-D4.	2198	100
617	AF116672	Homo sapiens	PRO1905	553	99
618	BC011707	Homo sapiens	nuclear receptor binding factor-2, clone MGC:19778 IMAGE:3687848, mRNA, complete cds.	1471	100
619	AL133606	Homo sapiens	cDNA DKFZp434A2017 (from clone DKFZp434A2017); partial cds.	5012	100
620	AAY58618	Homo sapiens	11-APR-2000 11-JUN-1999 Protein regulating gene expression PRGE-11.	1778	100
621	AF276707	Homo sapiens	carcinoma susceptibility protein (HCCA3) mRNA, complete cds.	1211	100
622	AF161554	Homo sapiens	HSPC069	3072	98
623	AAY73327	Homo sapiens	24-FEB-2000 04-MAY-1999 HTRM clone 052927 protein sequence.	1668	100
624	AAP60958	Homo sapiens	09-AUG-1991 20-JAN-1986 Sequence of human endogenous benzodiazepineoid(EBZD) polypeptide.	564	100
625	AK010262	Mus musculus	putative	1767	94
626	AK001317	Homo sapiens	FLJ10455 fis, clone NT2RP1001014.	2539	99
627	M80899	Homo sapiens	novel protein AHNAK mRNA, partial	6618	99

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			sequence.		
628	AAB21018	Homo sapiens	19-DEC-2000 28-JAN-2000 Human nucleic acid-binding protein, NuABP-22.	2629	99
629	L19183	Homo sapiens	MAC30 mRNA, 3' end.	901	97
630	BC000540	Homo sapiens	DKFZP434H132 protein, clone MGC:3034 IMAGE:3163610, mRNA, complete cds.	739	100
631	BC002857	Homo sapiens	clone MGC:3442 IMAGE:3636594, mRNA, complete cds.	1033	100
632	AAW78188	Homo sapiens	13-APR-1999 11-JUN-1998 Human secreted protein encoded by gene 63 clone HPMCC16.	1300	100
633	AK000587	Homo sapiens	FLJ20580 fis, clone REC00516.	848	100
634	AF116637	Homo sapiens	PRO1489	266	100
635	BC011495	Mus musculus	RIKEN cDNA 1110060O18 gene	1242	89
636	AL157473	Homo sapiens	cDNA DKFZp761L0424 (from clone DKFZp761L0424).	2160	99
637	AF005067	Homo sapiens	mRNA, complete cds.	1415	65
638	AL117532	Homo sapiens	cDNA DKFZp434E192 (from clone DKFZp434E192); partial cds.	3706	100
639	AF251441	Homo sapiens	motif and leucine zipper containing kinase AZK mRNA, complete cds.	4234	100
640	BC010493	Homo sapiens	clone MGC:16982 IMAGE:3048997, mRNA, complete cds.	2496	99
641	AK017531	Mus musculus	putative	795	50
642	BC000204	Homo sapiens	ribosomal protein S3A, clone MGC:3109 IMAGE:3350750, mRNA, complete cds.	1367	100
643	AF226076	Homo sapiens	(CHRAC15) mRNA, complete cds.	651	100
644	AK023267	Homo sapiens	FLJ13205 fis, clone NT2RP3004534, highly similar to Mouse oncogene (ect2) mRNA.	4129	99
645	AE006465	Homo sapiens	sequence section 4 of 8.	1605	100
646	AF272973	Homo sapiens	mRNA, complete cds.	1411	100
647	AF237982	Homo sapiens	7alpha-hydroxylase (CYP39A1) mRNA, complete cds.	2478	100
648	AK023139	Homo sapiens	FLJ13077 fis, clone NT2RP3001944, moderately similar to HYPOTHETICAL 47.6 KD PROTEIN C16C10.5 IN CHROMOSOME III.	1754	100
649	AF315591	Homo sapiens	2 (PUMH2) mRNA, complete cds.	2985	94
650	AF302691	Mus musculus	myelin expression factor-3-like protein	943	77
651	AL136592	Homo sapiens	cDNA DKFZp761I172 (from clone DKFZp761I172); complete cds.	1393	99
652	AAB58279	Homo sapiens	14-MAR-2001 08-MAR-2000 Lung cancer associated polypeptide sequence SEQ ID 617.	678	100
653	AF104927	Homo sapiens	ligase (TTLL1) mRNA, complete cds.	2260	100
654	AL163792	Arabidopsis thaliana	putative protein	587	49
655	AF233395	Homo sapiens	protein type 7 (SIRT7) mRNA, complete cds.	2086	100
656	AF233223	Homo sapiens	protein FBG2 (FBG2) mRNA, complete cds.	1602	100
657	AF317549	Homo sapiens	finger protein 268 (ZNF268) mRNA, complete cds.	2885	99
658	AK025426	Homo sapiens	FLJ21773 fis, clone COLF7849.	1172	100

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
659	AL049548	Homo sapiens	DNA sequence from clone 398G3 on chromosome 6q25.1-25.3. Contains the 3'part of the gene for the ortholog of rat CPG2, part of a novel gene, ESTs, STSs and GSSs, complete sequence.	771	100
660	AK000947	Homo sapiens	FLJ10085 fis, clone HEMBA1002161, moderately similar to MYOSIN HEAVY CHAIN, CARDIAC MUSCLE BETA ISOFORM.	931	100
661	AAY48487	Homo sapiens	08-DEC-1999 20-MAR-1998 Human breast tumour-associated protein 32.	432	36
662	AE003564	Drosophila melanogaster	CG13295 gene product	377	29
663	AB049955	Homo sapiens	mRNA for mitochondrial ribosomal protein S21, complete cds.	313	100
664	BC005357	Homo sapiens	Similar to RIKEN cDNA 1700073K01 gene, clone MGC:12458 IMAGE:3511019, mRNA, complete cds.	418	100
665	L08240	Homo sapiens	MG81 mRNA, partial cds.	3398	99
666	AK022732	Homo sapiens	FLJ12670 fis, clone NT2RM4002301.	1551	99
667	AAY57900	Homo sapiens	23-MAR-2000 28-MAY-1999 Human transmembrane protein HTMPN-24.	996	100
668	BC005805	Homo sapiens	clone MGC:1003 IMAGE:2988344, mRNA, complete cds.	862	100
669	AF151073	Homo sapiens	HSPC239	1535	100
670	AF151073	Homo sapiens	HSPC239	1209	100
671	AK000197	Homo sapiens	FLJ20190 fis, clone COLF0714.	1754	100
672	AJ010071	Homo sapiens	TOM1-like protein.	2444	99
673	AJ010071	Homo sapiens	TOM1-like protein.	1236	97
674	BC004395	Homo sapiens	Similar to apolipoprotein L, clone MGC:10978 IMAGE:3636011, mRNA, complete cds.	1700	100
675	AC016526	Homo sapiens	14 clone RP11-361H10 map 14q24.3, complete sequence.	2554	99
676	AJ279246	Homo sapiens	NPHS2 gene for podocin, exon 1 and joined CDS.	1939	100
677	AL136628	Homo sapiens	cDNA DKFZp564C182 (from clone DKFZp564C182); complete cds.	732	100
678	AF116636	Homo sapiens	PRO1488	362	100
679	AF116694	Homo sapiens	PRO2219	414	100
680	AK001867	Homo sapiens	FLJ11005 fis, clone PLACE1002996.	859	100
681	AK027746	Homo sapiens	FLJ14840 fis, clone OVARC1001916.	1531	99
682	AK026486	Homo sapiens	FLJ22833 fis, clone KAIA4266.	623	100
683	AK001421	Homo sapiens	FLJ10559 fis, clone NT2RP2002618, weakly similar to PROTEIN ARGININE N-METHYLTRANSFERASE 2 (EC 2.1.1.-).	1650	100
684	AK000521	Homo sapiens	FLJ20514 fis, clone KAT09756.	1313	100
685	X59869	Homo sapiens	TCF-1 mRNA for T cell factor 1 (splice form A).	1375	99
686	AK002135	Homo sapiens	FLJ11273 fis, clone PLACE1009338.	1419	100
687	AAY57896	Homo sapiens	23-MAR-2000 28-MAY-1999 Human transmembrane protein HTMPN-20.	733	100
688	AF189692	Homo sapiens	Cdc42 effector protein SPEC2 mRNA, complete cds.	452	100
689	AJ401269	Bos taurus	polyadenylate-binding protein 1	2439	99

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
690	AK016776	Mus musculus	putative	1801	69
691	AB038523	Homo sapiens	for MBIP, complete cds.	1772	100
692	AK000241	Homo sapiens	FLJ20234 fis, clone COLF5673.	2398	100
693	AJ245719	Homo sapiens	for brk kinase substrate (BKS gene).	2154	100
694	AAB97378	Homo sapiens	17-AUG-2001 08-NOV-2000 Human kringle domain containing protein 1.	1533	100
695	AB038523	Homo sapiens	for MBIP, complete cds.	1552	99
696	AK026105	Homo sapiens	FLJ22452 fis, clone HRC09667.	2419	100
697	AAB00187	Homo sapiens	08-FEB-2001 15-MAR-2000 Breast cancer protein BCN1.	635	45
698	AAY99425	Homo sapiens	08-AUG-2000 01-SEP-1999 Human PRO1558 (UNQ766) amino acid sequence SEQ ID NO:306.	1343	100
699	AK009886	Mus musculus	putative	1329	75
700	AK016154	Mus musculus	putative	1166	79
701	AF151072	Homo sapiens	HSPC238	843	100
702	AB045180	Homo sapiens	mRNA for toll-like receptor 9, complete cds.	5466	100
703	AAB58961	Homo sapiens	27-MAR-2001 08-MAR-2000 Breast and ovarian cancer associated antigen protein sequence SEQ ID 669.	460	98
704	BC007556	Homo sapiens	Similar to TEA domain family member 2, clone MGC:15481 IMAGE:2967735, mRNA, complete cds.	2365	100
705	AAB74726	Homo sapiens	12-JUN-2001 14-AUG-2000 Human membrane associated protein MEMAP-32.	1601	48
706	AF217413	Homo sapiens	3 isoform gene, complete cds, alternatively spliced.	4450	100
707	AAG01129	Homo sapiens	06-OCT-2000 21-FEB-2000 Human secreted protein, SEQ ID NO: 5210.	230	77
708	AK024066	Homo sapiens	FLJ14004 fis, clone Y79AA1002351.	1791	100
709	AF259799	Homo sapiens	acid transporter system A2 (ATA2) mRNA, complete cds.	2560	100
710	AJ250839	Homo sapiens	for serine/threonine protein kinase.	2227	100
711	AK027057	Homo sapiens	FLJ23404 fis, clone HEP18862.	410	91
712	AF116652	Homo sapiens	PRO0813	1023	100
713	AF208845	Homo sapiens	BM-003	861	65
714	AK001123	Homo sapiens	FLJ10261 fis, clone HEMBB1000975.	3127	100
715	BC002571	Homo sapiens	clone MGC:2481 IMAGE:3143135, mRNA, complete cds.	1419	99
716	BC004169	Homo sapiens	Similar to RIKEN cDNA 3110001A18 gene, clone MGC:2714 IMAGE:2821548, mRNA, complete cds.	1266	100
717	AK026147	Homo sapiens	FLJ22494 fis, clone HRC11131.	1152	99
718	AJ400877	Homo sapiens	gene, CEGP1 gene, C11orf14 gene, C11orf15 gene, C11orf16 gene and C11orf17 gene.	1094	99
719	AJ400877	Homo sapiens	gene, CEGP1 gene, C11orf14 gene, C11orf15 gene, C11orf16 gene and C11orf17 gene.	575	100
720	AJ400877	Homo sapiens	gene, CEGP1 gene, C11orf14 gene, C11orf15 gene, C11orf16 gene and C11orf17 gene.	379	97
721	AK021919	Homo sapiens	FLJ11857 fis, clone HEMBA1006807, moderately similar to Homo sapiens	1851	99

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			mRNA for SPOP.		
722	AAB93876	Homo sapiens	26-JUN-2001 28-JUL-2000 Human protein sequence SEQ ID NO:13784.	552	39
723	BC005827	Homo sapiens	H2B histone family, member Q, clone MGC:1729 IMAGE:2989788, mRNA, complete cds.	392	93
724	BC010929	Homo sapiens	clone MGC:13522 IMAGE:4291498, mRNA, complete cds.	932	99
725	BC010929	Homo sapiens	clone MGC:13522 IMAGE:4291498, mRNA, complete cds.	1446	100
726	AAU00875	Homo sapiens	04-JUL-2001 30-AUG-2000 Human cancer related protein 10.	2015	100
727	AK004371	Mus musculus	putative	1096	89
728	AK000846	Homo sapiens	FLJ20839 fis, clone ADKA02346.	1379	100
729	AF090939	Homo sapiens	HQ0641 PRO0641 mRNA, complete cds.	275	100
730	AF201950	Homo sapiens	protein mRNA, complete cds.	399	100
731	AF129756	Homo sapiens	gene, partial cds; and CLIC1, DDAH, G6b, G6c, G5b, G6d, G6e, G6f, BAT5, G5b, CSK2B, BAT4, G4, Apo M, BAT3, BAT2, AIF-1, 1C7, LST-1, LTB, TNF, and LTA genes, complete cds.	691	99
732	Z26593	Homo sapiens	rearranged mRNA for T-cell receptor alpha chain.	573	97
733	BC004366	Homo sapiens	clone MGC:10334 IMAGE:3641657, mRNA, complete cds.	1219	100
734	AK001243	Homo sapiens	FLJ10381 fis, clone NT2RM2002055.	2326	100
735	BC008947	Homo sapiens	Similar to RIKEN cDNA 1200008O12 gene, clone MGC:3422 IMAGE:3028566, mRNA, complete cds.	2319	99
736	AK000702	Homo sapiens	FLJ20695 fis, clone KAIA2502.	1554	100
737	AF090937	Homo sapiens	HQ0618 PRO0618 mRNA, complete cds.	492	100
738	BC005357	Homo sapiens	Similar to RIKEN cDNA 1700073K01 gene, clone MGC:12458 IMAGE:3511019, mRNA, complete cds.	597	99
739	BC005357	Homo sapiens	Similar to RIKEN cDNA 1700073K01 gene, clone MGC:12458 IMAGE:3511019, mRNA, complete cds.	603	99
740	AL390216	Homo sapiens	cDNA DKFZp762K222 (from clone DKFZp762K222).	1120	100
741	AK006724	Mus musculus	putative	1077	80
742	AK000615	Homo sapiens	FLJ20608 fis, clone KAT05987.	1038	98
743	AK000615	Homo sapiens	FLJ20608 fis, clone KAT05987.	778	98
744	AAY99669	Homo sapiens	03-NOV-2000 23-NOV-1999 Human GTPase associated protein-20.	1013	99
745	AL137516	Homo sapiens	cDNA DKFZp564M2178 (from clone DKFZp564M2178); partial cds.	3506	99
746	BC006006	Homo sapiens	Similar to RIKEN cDNA 1810046J19 gene, clone MGC:14832 IMAGE:4283597, mRNA, complete cds.	597	100
747	AB002819	Perilla frutescens	actin	142	96
748	AJ271290	Homo sapiens	for facilitative glucose transporter GLUT11 (SLC2A11 gene).	866	99
749	AK000157	Homo sapiens	FLJ20150 fis, clone COL08263.	1559	99

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
750	AF231410	Homo sapiens	sperm protein ropporin mRNA, complete cds.	205	87
751	AB044755	Homo sapiens	mRNA for basic-helix-loop-helix protein, complete cds.	1723	100
752	AK001610	Homo sapiens	FLJ10748 fis, clone NT2RP3001819, weakly similar to RING CANAL PROTEIN.	1852	100
753	AF208864	Homo sapiens	ARF	688	99
754	AF209931	Homo sapiens	protein mRNA, partial cds.	1222	96
755	AF064604	Homo sapiens	protein mRNA, partial cds.	1350	93
756	AK000042	Homo sapiens	FLJ20035 fis, clone COL00213.	2029	100
757	AF208864	Homo sapiens	ARF	688	99
758	AK016624	Mus musculus	putative	847	84
759	AF332890	Homo sapiens	zinc finger FEZL	1561	99
760	AK000602	Homo sapiens	FLJ20595 fis, clone KAT08558.	764	100
761	AAB73227	Homo sapiens	11-MAY-2001 11-AUG-2000 Human phosphatase NP_060232_h.	2733	99
762	AF016903	Homo sapiens	precursor mRNA, partial cds.	8478	100
763	BC002912	Homo sapiens	clone MGC:11279 IMAGE:3944940, mRNA, complete cds.	1512	98
764	AB035179	Homo sapiens	for HES6, complete cds.	1143	98
765	AK000506	Homo sapiens	FLJ20499 fis, clone KAT09034.	3811	99
766	AAG71494	Homo sapiens	31-JUL-2001 06-OCT-2000 Human olfactory receptor polypeptide, SEQ ID NO: 1175.	613	98
767	BC001005	Homo sapiens	cytochrome c oxidase subunit VIIc, clone MGC:8432 IMAGE:2821167, mRNA, complete cds.	329	100
768	AF104260	Homo sapiens	mRNA, partial cds.	1327	51
769	AF116709	Homo sapiens	PRO2605	642	100
770	AF176330	Homo sapiens	(PCBP4) mRNA, complete cds.	2041	100
771	AF169226	Homo sapiens	conserved domain protein 1 (ACDP1) mRNA, complete cds.	972	100
772	Z83851	Homo sapiens	DNA sequence from clone 989H11 on chromosome 22q13.1-13.2. Contains part of a novel gene, ESTs, GSSs and four putative CpG islands, complete sequence.	474	100
773	AK000130	Homo sapiens	FLJ20123 fis, clone COL06041.	998	100
774	AC004908	Homo sapiens	clone RP5-855D21, complete sequence.	171	91
775	X68879	Homo sapiens	EMX1 mRNA.	819	100
776	AAR78692	Homo sapiens	15-MAR-1996 24-DEC-1993 Human skeletal muscle stress protein, p20.	832	100
777	AF042831	Homo sapiens	transcription factor FREAC-10 (FKHL18) mRNA, partial cds.	615	100
778	AK001050	Homo sapiens	FLJ10188 fis, clone HEMBA1004693.	1312	100
779	AL096817	Homo sapiens	DNA sequence from clone RP1-102H19 on chromosome 6q15-16.1. Contains an HSP60 (TCP-1/cpn60 chaperonin family) pseudogene, three novel genes, ESTs, STSs and GSSs, complete sequence.	320	100
781	AF202922	Homo sapiens	(LRP16) mRNA, complete cds.	1511	95
782	AL365514	Homo sapiens	human gene mapping to chomosome 22.	2255	100
783	AF023859	Papio hamadryas	cyclophilin A	538	95
784	X03491	Mus musculus	57 kd keratin (aa 1-524)	2099	80

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
785	AAB27242	Homo sapiens	27-MAR-2001 10-MAY-2000 Human EXMAD-20 SEQ ID NO: 20.	2327	98
786	AAY99414	Homo sapiens	08-AUG-2000 01-SEP-1999 Human PRO1461 (UNQ742) amino acid sequence SEQ ID NO:269.	2270	100
787	AF043350	Homo sapiens	protein 1 (LSP1) gene, LSP1-5 allele, partial cds.	361	100
788	BC011551	Homo sapiens	clone MGC:19971 IMAGE:4561164, mRNA, complete cds.	1606	88
789	AL050256	Homo sapiens	human gene mapping to chomosome 22.	881	100
790	AY014302	Homo sapiens	gene, exon 2 and complete cds.	1409	100
791	AK000314	Homo sapiens	FLJ20307 fis, clone HEP07254.	5380	99
792	AK023886	Homo sapiens	FLJ13824 fis, clone THYRO1000505.	1377	100
793	AK019547	Mus musculus	putative	265	96
794	AK005789	Mus musculus	putative	475	97
795	AK001783	Homo sapiens	FLJ10921 fis, clone OVARC1000411.	1246	100
796	AK027598	Homo sapiens	FLJ14692 fis, clone NT2RP2005344, weakly similar to PROBABLE CALCIUM-TRANSPORTING ATPASE 5 (EC 3.6.1.38).	3134	99
797	U60269	Homo sapiens	endogenous retrovirus HERV-K(HML6) proviral clone HML6.17 putative polymerase and envelope genes, partial cds, and 3'LTR.	381	100
798	AL137651	Homo sapiens	cDNA DKFZp434O0213 (from clone DKFZp434O0213); partial cds.	1366	100
799	AK000061	Homo sapiens	FLJ20054 fis, clone COL00849.	1751	99
800	AF233588	Homo sapiens	(RIS) mRNA, complete cds.	1353	100
801	S76838	Mus sp.	Dbs	1469	49
802	AB033168	Mus musculus	nuclear protein ZAP	1946	89
803	AB049591	Homo sapiens	related with psoriasis, complete cds.	647	100
804	AF093249	Homo sapiens	isoform 4 (PHRET1) mRNA, alternatively spliced, complete cds.	1046	100
805	AL049679	Homo sapiens	gene from PAC 97K10, chromosome X, similar to heparan-sulphate 6-sulfotransferase.	1527	100
806	AB015329	Homo sapiens	mRNA, partial cds.	1055	97
807	AF077034	Homo sapiens	HSPC010	163	96
808	AF241833	Mus musculus	secretory carrier membrane protein 5	1256	98
809	AK001352	Homo sapiens	FLJ10490 fis, clone NT2RP2000233.	697	100
810	AF138860	Homo sapiens	PRO0843	649	100
811	Z72496	Homo sapiens	MUC5B gene (partial).	18275	100
812	AK000361	Homo sapiens	FLJ20354 fis, clone HEP15013.	3585	99
813	AK001072	Homo sapiens	FLJ10210 fis, clone HEMBA1006344, weakly similar to RADIXIN.	2372	100
814	AK001707	Homo sapiens	FLJ10845 fis, clone NT2RP4001372, weakly similar to IRREGULAR CHIASM C-ROUGHEST PROTEIN PRECURSOR.	2161	100
815	S79854	Homo sapiens	3 iodothyronine deiodinase mRNA, complete cds.	774	100
816	AB036704	Homo sapiens	mRNA for phosphodiesterase 11A, complete cds.	2541	100
817	BC010181	Homo sapiens	clone MGC:20197 IMAGE:4543414, mRNA, complete cds.	387	89
818	AAB68074	Homo sapiens	09-JUL-2001 10-NOV-2000 Amino acid	1960	99

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			sequence of a human chordin-like homologue splice variant.		
819	AF227516	Homo sapiens	mRNA, complete cds.	1444	97
820	AF077202	Homo sapiens	HSPC016	100	100
821	AK002945	Mus musculus	putative	615	94
822	AK000047	Homo sapiens	FLJ20040 fis, clone COL00417.	193	97
823	AF119878	Homo sapiens	PRO2353	401	100
824	BC005827	Homo sapiens	H2B histone family, member Q, clone MGC:1729 IMAGE:2989788, mRNA, complete cds.	385	100
825	AAB73230	Homo sapiens	11-MAY-2001 11-AUG-2000 Human phosphatase AA493915_h.	423	97
826	AK000513	Homo sapiens	FLJ20506 fis, clone KAT09493.	707	100
827	AK001021	Homo sapiens	FLJ10159 fis, clone HEMBA1003528.	1471	100
828	AJ002535	Homo sapiens	for obscurin (OBSCN gene).	466	100
829	AJ243662	Homo sapiens	for NICE-1 protein.	566	100
830	AK000268	Homo sapiens	FLJ20261 fis, clone COLF7630.	2659	100
831	AC005396	Arabidopsis thaliana	putative proline-rich protein	113	32
832	AJ249977	Homo sapiens	for AMP-activated protein kinase gamma 3 subunit (AMPK gamma 3 gene).	2518	99
833	AAY44985	Homo sapiens	23-MAY-2000 27-JUL-1999 Human epidermal protein-2.	616	88
834	AJ006692	Homo sapiens	KerB gene.	923	76
835	M88166	Sus scrofa	small proline-rich protein	190	59
836	AK000139	Homo sapiens	FLJ20132 fis, clone COL06441.	2479	100
837	AK000520	Homo sapiens	FLJ20513 fis, clone KAT09741.	805	99
838	AC004744	Homo sapiens	clone GS1-465N13 from 7p15-p21, complete sequence.	293	98
839	AJ002535	Homo sapiens	for obscurin (OBSCN gene).	466	100
840	AAY87354	Homo sapiens	11-MAY-2000 25-JUN-1999 Human signal peptide containing protein HSPP-131 SEQ ID NO:131.	1546	100
841	AK000054	Homo sapiens	FLJ20047 fis, clone COL00577.	4964	100
842	AAB47129	Homo sapiens	04-JUN-2001 14-SEP-2000 CDIFF-7, Incyte ID No. 2027937CD1.	672	100
843	Z81024	Homo sapiens	mRNA for TCR alpha (TCRAV).	604	90
844	AK001720	Homo sapiens	FLJ10858 fis, clone NT2RP4001555.	3226	99
845	AE000660	Homo sapiens	receptor alpha delta locus from bases 501613 to 752736 (section 3 of 5) of the Complete Nucleotide Sequence.	561	99
846	U61084	Homo sapiens	protein mRNA, complete cds.	1281	97
847	AF161550	Homo sapiens	HSPC065	954	99
848	AK001002	Homo sapiens	FLJ10140 fis, clone HEMBA1003179, moderately similar to PROBABLE TRNA (5-METHYLAMINOMETHYL-2-THIOURIDYLATE)-METH YLTRANSFERASE (EC 2.1.1.61).	896	99
849	AJ406946	Homo sapiens	for keratin associated protein 9.2 (KRTAP9.2 gene).	1079	95
850	AF339106	Mus musculus	forkhead-related transcription factor 2	1480	99
851	AF081797	Mus musculus	high cysteine keratin-associated protein 12.1	469	58
852	AF071081	Mycobacterium tuberculosis	proline-rich mucin homolog	121	36

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
853	AF116686	Homo sapiens	PRO2116	192	100
854	AF070655	Homo sapiens	F1F0-type ATP synthase subunit g	443	89
855	AAY41710	Homo sapiens	07-DEC-1999 08-MAR-1999 Human PRO618 protein sequence.	4232	98
856	AAB47276	Homo sapiens	06-AUG-2001 12-JUL-2000 hOAT5.	887	98
857	AF113013	Homo sapiens	PRO0806	345	100
858	X60661	Rattus rattus	potential ligand-binding protein	344	74
859	AF119902	Homo sapiens	PRO2832	406	100
860	AK009462	Mus musculus	putative	1723	100
861	AAB95296	Homo sapiens	26-JUN-2001 28-JUL-2000 Human protein sequence SEQ ID NO:17523.	4692	99
862	AB017927	Homo sapiens	mRNA for p53DINP1b, complete cds.	878	100
863	AAB83845	Homo sapiens	23-JUL-2001 30-OCT-2000 Amino acid sequence of a human protein expressed in tumour cells.	1346	54
864	AX149579	Homo sapiens	DNA encoding a transmembrane serine protease (Endotheliase 2-S) protein	562	98
865	BC012048	Homo sapiens	clone IMAGE:3502817, mRNA, partial cds.	1225	99
866	AK000575	Homo sapiens	FLJ20568 fis, clone REC00775.	664	99
867	X76383	Homo sapiens	mRNA for HE3(alpha).	807	100
868	AF286598	Homo sapiens	mRNA, complete cds.	2381	100
869	AK022643	Homo sapiens	FLJ12581 fis, clone NT2RM4001140, weakly similar to HOMEOBOX PROTEIN MSH-D.	721	92
870	AF119891	Homo sapiens	PRO2706	363	100
871	AK009258	Mus musculus	putatiye	1246	80
872	U66412	Mus musculus	adenomatous polyposis coli	133	88
873	AK001162	Homo sapiens	FLJ10300 fis, clone NT2RM2000030.	184	100
874	AL033518	Homo sapiens	DNA sequence from clone RP3-322I12 on chromosome 6p21.1-21.31. Contains part of the gene for a novel protein similar to C. elegans C05C8.6 (Tr:016313), STSs and GSSs, complete sequence.	199	100
875	AF116601	Homo sapiens	PRO0128	446	100
876	AF156889	Homo sapiens	homeobox protein 3 isoform b (LHX3) mRNA, complete cds.	2148	100
877	AK026671	Homo sapiens	FLJ23018 fis, clone LNG00903.	385	100
878	AAY92515	Homo sapiens	10-AUG-2000 06-OCT-1999 Human OXRE-12.	2523	99
879	AL136818	Homo sapiens	cDNA DKFZp434F1726 (from clone DKFZp434F1726).	1736	99
880	AB055311	Homo sapiens	for RanBPM, complete cds.	2172	67
881	AF006465	Mus musculus	B cell antigen receptor Ig beta associated protein 1	1286	61
882	AF143956	Mus musculus	coronin-2	1020	72
883	AK008237	Mus musculus	putative	653	84
884	AK008237	Mus musculus	putative	653	84
885	AF221846	Homo sapiens	gastric protein ZG12P mRNA, complete cds.	182	100
886	BC001005	Homo sapiens	cytochrome c oxidase subunit VIIc, clone MGC:8432 IMAGE:2821167, mRNA, complete cds.	304	93
887	X01715	Homo sapiens	gene fragment for the acetylcholine	2543	100

SEQ ID NO:	Accession No.	Species	Description	Score	% Identity
			receptor gamma subunit precursor (exons 1 and 2).		
888	AK001974	Homo sapiens	FLJ11112 fis, clone PLACE1005925.	955	100
889	AF212016	Homo sapiens	receptor 9 (IL1R9) mRNA, complete cds.	3607	100
890	D88437	Homo sapiens	for G-protein coupled receptor SALPR, complete cds.	2455	100
891	AK002298	Mus musculus	putative	833	97
892	X96389	Bos taurus	procollagen I N-proteinase	440	34

TABLE 3

SEQ ID NO:	Accession No.	Description	Results*
451	PD01719	PRECURSOR GLYCOPROTEIN SIGNAL RE.	PD01719A 12.89 8.200e-17 343-371
452	BL00388	Proteasome A-type subunits proteins.	BL00388A 23.14 5.875e-40 5-51 BL00388B 31.38 6.538e-29 64-106 BL00388D 20.71 1.391e-26 147-178 BL00388C 18.79 2.000e-22 119-141
453	BL00388	Proteasome A-type subunits proteins.	BL00388B 31.38 6.538e-29 33-75 BL00388D 20.71 1.391e-26 116-147 BL00388C 18.79 2.000e-22 88-110
454	BL00064	L-lactate dehydrogenase proteins.	BL00064C 17.28 8.442e-22 293-338 BL00064A 21.16 5.574e-12 184-222
459	BL01187	Calcium-binding EGF-like domain proteins pattern proteins.	BL01187B 12.04 1.257e-10 218-234
460	BL00107	Protein kinases ATP-binding region proteins.	BL00107B 13.31 9.100e-15 199-215
462	PR00678	PI3 KINASE P85 REGULATORY SUBUNIT SIGNATURE	PR00678H 9.13 1.529e-11 64-87
463	PR00678	PI3 KINASE P85 REGULATORY SUBUNIT SIGNATURE	PR00678H 9.13 1.529e-11 64-87
469	PD00930	PROTEIN GTPASE DOMAIN ACTIVATION.	PD00930B 33.72 6.250e-17 446-487 PD00930A 25.62 2.841e-13 343-369
472	PR00302	LUPUS LA PROTEIN SIGNATURE	PR00302A 11.32 3.318e-14 222-240
473	PF00777	Sialyltransferase family.	PF00777C 18.60 9.416e-26 363-418 PF00777D 22.05 3.681e-11 511-557
476	BL00360	Ribosomal protein S9 proteins.	BL00360B 20.22 5.705e-19 317-353 BL00360C 17.65 4.857e-18 370-397
479	BL00057	Ribosomal protein S18 proteins.	BL00057 24.94 8.800e-14 81-129
482	BL00299	Ubiquitin domain proteins.	BL00299 28.84 1.000e-40 16-68 BL00299 28.84 1.000e-40 92-144
483	BL00039	DEAD-box subfamily ATP-dependent helicases proteins.	BL00039D 21.67 9.000e-37 321-367 BL00039A 18.44 3.893e-24 28-67 BL00039C 15.63 8.269e-17 165-189 BL00039B 19.19 4.818e-14 73-99
485	PR00828	FORMIN SIGNATURE	PR00828B 5.23 8.218e-10 382-405
489	PR00581	PROSTANOID EP2 RECEPTOR SIGNATURE	PR00581E 3.48 9.875e-10 4-20
490	BL00615	C-type lectin domain proteins.	BL00615A 16.68 8.200e-11 113-131
492	BL00039	DEAD-box subfamily ATP-dependent helicases proteins.	BL00039D 21.67 4.176e-23 391-437 BL00039A 18.44 7.065e-16 118-157 BL00039B 19.19 5.395e-12 158-184 BL00039C 15.63 9.820e-11 241-265
493	BL00479	Phorbol esters / diacylglycerol binding domain proteins.	BL00479B 12.57 9.518e-09 472-488

SEQ ID NO:	Accession No.	Description	Results*
494	PR00929	AT-HOOK-LIKE DOMAIN SIGNATURE	PR00929B 4.38 1.000e-10 49-61
499	BL00678	Trp-Asp (WD) repeat proteins proteins.	BL00678 9.67 4.000e-10 339-350
500	PF00646	F-box domain proteins.	PF00646A 14.37 4.375e-09 74-88
503	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 3.800e-30 6-45
504	DM00191	w SPAC8A4.04C RESISTANCE SPAC8A4.05C DAUNORUBICIN.	DM00191A 8.16 4.360e-09 210-223
506	PR00060	RIBOSOMAL PROTEIN L16 SIGNATURE	PR00060A 10.94 6.023e-09 117-130
507	PF00646	F-box domain proteins.	PF00646A 14.37 9.036e-10 13-27
508	BL00632	Ribosomal protein S4 proteins.	BL00632 23.79 2.821e-12 104-147
510	BL01191	Ribosomal protein S3Ae proteins.	BL01191A 15.57 1.000e-40 13-64 BL01191B 13.33 1.000e-40 89-140
512	PD00930	PROTEIN GTPASE DOMAIN ACTIVATION.	PD00930B 33.72 6.063e-25 162-203 PD00930A 25.62 8.297e-15 41-67
513	BL00604	Synaptophysin / synaptoporin proteins.	BL00604F 5.96 7.718e-10 567-612
515	BL01152	Hypothetical hesB/yadR/yfhF family proteins.	BL01152C 25.93 1.900e-29 81-128 BL01152B 20.12 6.121e-11 48-74
516	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 1.000e-09 27-42
518	BL00218	Amino acid permeases proteins.	BL00218D 21.49 3.797e-11 243-288 BL00218B 21.44 1.621e-10 75-107 BL00218E 23.30 3.520e-10 324-364
523	PR00836	SOMATOTROPIN HORMONE FAMILY SIGNATURE	PR00836B 16.59 2.895e-16 101-120 PR00836D 13.05 1.621e-13 195-210 PR00836A 14.40 2.800e-13 79-93 PR00836C 11.95 4.913e-13 179-196
526	BL00164	Enolase proteins.	BL00164B 16.22 1.000e-40 98-141 BL00164C 15.66 1.000e-40 144-194 BL00164G 12.13 1.000e-40 380-419 BL00164F 10.48 3.813e-39 313-349 BL00164D 21.97 2.588e-38 220-263 BL00164A 11.58 1.529e-27 32-55 BL00164E 8.80 9.100e-20 287-302
529	BL00790	Receptor tyrosine kinase class V proteins.	BL00790R 16.20 3.516e-09 21-65
530	PR00288	PUROTHIONIN SIGNATURE	PR00288B 13.09 9.870e-09 3-17
536	DM00250	kw ANNEXIN ANTIGEN PROLINE TUMOR.	DM00250B 13.84 8.541e-09 426-450
540	BL00495	Apple domain proteins.	BL00495G 12.47 8.920e-09 80-109
542	PR00080	ALCOHOL DEHYDROGENASE SUPERFAMILY SIGNATURE	PR00080C 17.16 4.750e-12 147-167
551	PR00926	MITOCHONDRIAL CARRIER PROTEIN SIGNATURE	PR00926F 17.75 1.964e-20 4-27

SEQ ID NO:	Accession No.	Description	Results*
552	BL00795	Involucrin proteins.	BL00795C 17.06 2.286e-12 103-148 BL00795C 17.06 5.208e-12 102-147 BL00795C 17.06 8.953e-10 99-144 BL00795C 17.06 1.000e-09 114-159 BL00795C 17.06 1.400e-09 97-142 BL00795C 17.06 3.200e-09 104-149 BL00795C 17.06 4.100e-09 101-146 BL00795C 17.06 4.800e-09 100-145
556	PF00628	PHD-finger.	PF00628 15.84 6.806e-09 77-92
559	PR00041	CAMP RESPONSE ELEMENT BINDING (CREB) PROTEIN SIGNATURE	PR00041E 7.20 7.072e-12 219-240
564	BL01119	Copper-fist domain proteins.	BL01119B 18.30 2.385e-09 3818-3836
568	BL00814	Adrenodoxin family, iron-sulfur binding region proteins.	BL00814B 23.55 9.372e-22 127-165 BL00814A 15.33 3.769e-15 100-118
570	PF00152	tRNA synthetases class II.	PF00152D 21.30 4.774e-29 434-473 PF00152C 28.03 7.107e-25 110-147
571	PR00608	CLASS II CYTOCHROME C SIGNATURE	PR00608A 13.74 7.000e-09 78-102
574	BL00376	S-adenosylmethionine synthetase proteins.	BL00376A 10.62 1.000e-40 19-74 BL00376D 18.36 1.000e-40 157-201 BL00376C 11.94 3.571e-38 122-157 BL00376B 14.91 3.500e-19 99-116
579	BL00415	Synapsins proteins.	BL00415N 4.29 6.058e-12 328-372
580	BL00475	Ribosomal protein L15 proteins.	BL00475B 8.20 6.769e-09 46-56 BL00475D 16.25 9.578e-09 151-173
581	BL00678	Trp-Asp (WD) repeat proteins proteins.	BL00678 9.67 4.000e-13 241-252
583	PD02784	PROTEIN NUCLEAR RIBONUCLEOPROTEIN.	PD02784B 26.46 3.629e-13 96-139 PD02784C 20.76 6.894e-09 228-274
584	BL00417	Synaptobrevin proteins.	BL00417B 18.48 1.000e-40 59-113 BL00417A 7.74 3.700e-34 31-59
585	BL01013	Oxysterol-binding protein family proteins.	BL01013D 26.81 9.578e-17 267-311 BL01013C 9.97 6.308e-13 91-101 BL01013B 11.33 3.717e-12 65-76
586	PR00302	LUPUS LA PROTEIN SIGNATURE	PR00302A 11.32 3.647e-13 99-117
589	BL00018	EF-hand calcium-binding domain proteins.	BL00018 7.41 1.000e-12 209-222
590	BL00708	Prolyl endopeptidase family serine proteins.	BL00708B 24.91 2.235e-15 619-650
593	BL01032	Protein phosphatase 2C proteins.	BL01032H 11.25 1.000e-10 446-459 BL01032C 6.14 4.474e-09 175-185
596	BL01115	GTP-binding nuclear protein ran	BL01115A 10.22 3.600e-16 8-52

SEQ ID NO:	Accession No.	Description	Results*
		proteins.	
597	BL00226	Intermediate filaments proteins.	BL00226D 19.10 4.450e-18 113-160
598	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 2.800e-14 139-152 PD00066 13.92 2.800e-14 195-208 PD00066 13.92 5.200e-14 167-180 PD00066 13.92 5.500e-13 363-376 PD00066 13.92 1.857e-12 223-236 PD00066 13.92 2.714e-12 419-432 PD00066 13.92 9.143e-12 279-292 PD00066 13.92 9.143e-12 307-320 PD00066 13.92 4.913e-11 251-264 PD00066 13.92 1.346e-10 335-348 PD00066 13.92 2.200e-09 391-404
599	BL00194	Thioredoxin family proteins.	BL00194 12.16 5.500e-14 176-189 BL00194 12.16 4.913e-13 64-77
604	PD00289	PROTEIN SH3 DOMAIN REPEAT PRESYNA.	PD00289 9.97 9.550e-11 62-76
607	BL00960	BTG1 family proteins.	BL00960C 12.68 3.647e-26 23-45
609	PR00366	ENDOTHELIN RECEPTOR SIGNATURE	PR00366A 14.10 4.222e-09 5-25
611	BL00383	Tyrosine specific protein phosphatases proteins.	BL00383E 10.35 6.368e-09 93-104
612	BL00290	Immunoglobulins and major histocompatibility complex proteins.	BL00290B 13.17 8.773e-10 266-284
614	BL00415	Synapsins proteins.	BL00415C 7.09 3.182e-09 415-445
616	PF00628	PHD-finger.	PF00628 15.84 5.125e-11 451-466
619	BL00322	Histone H3 proteins.	BL00322B 13.68 8.514e-10 933-986
622	PD02411	PROTEIN TRANSCRIPTION REGULATION NUCLEAR.	PD02411 21.89 4.214e-15 183-217
624	BL00880	Acyl-CoA-binding protein.	BL00880 17.52 1.000e-40 96-146
628	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 7.300e-17 383-396 PD00066 13.92 3.400e-14 439-452 PD00066 13.92 4.000e-14 355-368 PD00066 13.92 8.000e-13 327-340 PD00066 13.92 9.500e-13 411-424
631	BL00226	Intermediate filaments proteins.	BL00226D 19.10 4.667e-11 121-168
632	PD01613	RIBOSOME FACTOR PROTEIN RECYCLIN.	PD01613 23.39 6.121e-17 169-215
636	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 9.500e-10 842-857
637	PF00855	PWWP domain proteins.	PF00855 13.75 3.872e-17 1078-1095
638	PR00671	INHIBIN BETA B CHAIN SIGNATURE	PR00671C 4.18 9.671e-10 549-569
639	BL00240	Receptor tyrosine kinase class III proteins.	BL00240F 17.74 7.645e-11 157-205 BL00240G 28.45 1.818e-10 204-257
642	BL01191	Ribosomal protein S3Ae proteins.	BL01191A 15.57 1.000e-40 13-64 BL01191B 13.33 1.000e-40 89-140 BL01191C 16.50 1.000e-40 180-232

SEQ ID NO:	Accession No.	Description	Results*
643	PR00950	FLAGELLAR BIOSYNTHETIC PROTEIN FLHB SIGNATURE	PR00950B 14.12 6.571e-09 92-115
644	BL00741	Guanine-nucleotide dissociation stimulators CDC24 family sign.	BL00741B 14.27 7.808e-09 558-581
646	PD02059	CORE POLYPROTEIN PROTEIN GAG CONTAINS: P.	PD02059B 24.48 7.211e-09 125-160
647	BL00086	Cytochrome P450 cysteine heme-iron ligand proteins.	BL00086 20.87 7.395e-13 404-436
649	PF00806	Pumilio-family RNA binding domain proteins (aka PUM-HD, Pumilio homol.	PF00806B 11.32 4.176e-12 766-776 PF00806C 7.81 5.263e-11 838-847 PF00806C 7.81 7.632e-09 694-703
650	PR00221	CAULIMOVIRUS COAT PROTEIN SIGNATURE	PR00221H 12.82 7.614e-09 298-312
651	PF00023	Ank repeat proteins.	PF00023A 16.03 9.571e-11 50-66
654	BL01279	Protein-L-isoaspartate(D-aspartate) O-methyltransferase signa.	BL01279A 24.27 6.967e-10 90-138
657	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 1.000e-14 351-368 BL00028 16.07 4.706e-14 267-284 BL00028 16.07 7.882e-14 71-88 BL00028 16.07 5.500e-13 183-200 BL00028 16.07 5.950e-13 127-144 BL00028 16.07 2.174e-12 491-508 BL00028 16.07 2.957e-12 323-340 BL00028 16.07 8.043e-12 463-480 BL00028 16.07 9.217e-12 435-452 BL00028 16.07 2.038e-11 211-228 BL00028 16.07 3.769e-11 15-32 BL00028 16.07 4.115e-11 379-396 BL00028 16.07 8.615e-11 295-312 BL00028 16.07 8.962e-11 99-116 BL00028 16.07 5.200e-10 155-172 BL00028 16.07 9.100e-10 43-60 BL00028 16.07 9.100e-10 239-256
658	PF00850	Histone deacetylase family.	PF00850E 8.88 4.750e-12 52-78 PF00850D 14.76 8.696e-11 17-41 PF00850G 22.75 5.382e-10 115-157
660	PR00193	MYOSIN HEAVY CHAIN SIGNATURE	PR00193A 15.41 6.294e-22 114-134
661	BL00478	LIM domain proteins.	BL00478B 14.79 5.500e-13 11-26
665	PF00566	Probable rabGAP domain proteins.	PF00566B 11.92 6.100e-09 330-336
676	BL01270	Band 7 protein family proteins.	BL01270D 20.87 1.509e-21 232-270 BL01270B 18.74 4.136e-16 164-203 BL01270A 9.40 8.953e-13 124-137 BL01270E 13.03 8.500e-12 270-299
685	PD02448	TRANSCRIPTION PROTEIN DNA-BINDIN.	PD02448A 9.37 3.927e-09 159-198
686	PR00625	DNAJ PROTEIN FAMILY SIGNATURE	PR00625D 11.93 7.828e-10 61-72
689	PF00658	Poly-adenylate binding protein, unique domain proteins.	PF00658B 28.57 1.000e-40 105-152 PF00658C 16.33 8.500e-36 421-458
696	PF00566	Probable rabGAP domain proteins.	PF00566A 12.64 1.409e-11 210-

SEQ ID NO:	Accession No.	Description	Results*
			220
698	BL01100	NNMT/PNMT/TEMT family of methyltransferases proteins.	BL01100E 12.25 9.277e-09 171-215
699	BL00569	Myelin basic protein.	BL00569A 16.70 3.632e-09 147-190
702	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019A 11.19 7.261e-10 679-693 PR00019B 11.36 7.300e-10 676-690 PR00019B 11.36 8.650e-10 520-534 PR00019B 11.36 4.240e-09 122-136 PR00019B 11.36 4.240e-09 307-321 PR00019A 11.19 4.333e-09 417-431 PR00019A 11.19 8.000e-09 222-236
703	BL00025	P-type 'Trefoil' domain proteins.	BL00025 17.17 9.217e-21 53-74
704	BL00554	TEA domain proteins.	BL00554A 11.66 1.000e-40 62-107 BL00554C 12.10 1.000e-40 326-379 BL00554D 12.30 1.000e-40 389-444 BL00554B 10.31 8.875e-39 262-303
706	PR00878	CHOLINESTERASE SIGNATURE	PR00878F 5.37 4.780e-13 503-516
709	BL00594	Aromatic amino acids permeases proteins.	BL00594A 16.75 5.688e-10 76-120
710	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 3.647e-20 136-167 BL00107B 13.31 6.727e-13 205-221
711	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 3.250e-35 14-53
713	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 4.857e-09 6-23
715	PR00111	ALPHA/BETA HYDROLASE FOLD SIGNATURE	PR00111A 11.49 4.200e-11 123-139
721	PF00651	BTB (also known as BR-C/Ttk) domain proteins.	PF00651 15.00 2.895e-11 213-226
722	BL00069	Glucose-6-phosphate dehydrogenase proteins.	BL00069C 16.11 7.723e-09 19-50
723	PR00621	HISTONE H2B SIGNATURE	PR00621A 12.25 8.714e-23 38-57 PR00621B 4.91 5.034e-21 57-78
724	BL00919	Deoxyribonuclease I proteins.	BL00919F 14.41 9.010e-09 108-143
725	BL00919	Deoxyribonuclease I proteins.	BL00919F 14.41 9.010e-09 108-143
726	BL00790	Receptor tyrosine kinase class V proteins.	BL00790I 20.01 2.375e-12 192-223
727	BL01115	GTP-binding nuclear protein ran proteins.	BL01115A 10.22 3.089e-10 23-67
731	BL00983	Ly-6 / u-PAR domain proteins.	BL00983C 12.69 4.981e-09 83-99
732	DM00031	IMMUNOGLOBULIN V REGION.	DM00031B 15.41 1.797e-09 79-113
740	PF00078	Reverse transcriptase (RNA-dependent DNA polymerase).	PF00078A 8.82 9.438e-09 803-811
744	BL01020	SAR1 family proteins.	BL01020C 15.35 7.038e-20 71-122
745	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 3.000e-13 727-740 PD00066 13.92 1.000e-12 671-684

SEQ ID NO:	Accession No.	Description	Results*
			PD00066 13.92 5.286e-12 699-712 PD00066 13.92 6.143e-12 428-441 PD00066 13.92 9.571e-12 456-469
747	PR00190	ACTIN SIGNATURE	PR00190F 7.80 7.506e-09 33-53
748	BL00216	Sugar transport proteins.	BL00216B 27.64 4.512e-16 127-177
749	PF00622	Domain in SPla and the RYanodine Receptor.	PF00622B 21.00 9.795e-09 166-188
750	DM01513	CAMP-DEPENDENT PROTEIN KINASE REGULATORY CHAIN.	DM01513A 13.61 1.491e-09 10-51
751	BL00038	Myc-type, 'helix-loop-helix' dimerization domain proteins.	BL00038B 16.97 4.750e-14 84-105 BL00038A 13.61 4.750e-11 57-73
753	BL01019	ADP-ribosylation factors family proteins.	BL01019A 13.20 4.882e-24 47-87
754	PD00930	PROTEIN GTPASE DOMAIN ACTIVATION.	PD00930B 33.72 7.000e-17 23-64
757	BL01019	ADP-ribosylation factors family proteins.	BL01019A 13.20 4.882e-24 47-87
759	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 6.400e-13 279-296
761	DM01724	kw ALLERGEN POLLEN CIMI HOL-LI.	DM01724 8.14 9.526e-09 192-212
762	DM00758	AGRIN.	DM00758 13.12 8.250e-14 341-357
763	PR00421	THIOREDOXIN FAMILY SIGNATURE	PR00421B 11.40 7.400e-09 29-39
764	BL00038	Myc-type, 'helix-loop-helix' dimerization domain proteins.	BL00038A 13.61 5.667e-10 34-50
766	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245A 18.03 2.373e-13 26-48
770	PF00013	KH domain proteins family of RNA binding proteins.	PF00013 5.78 7.300e-09 32-44
775	BL00027	'Homeobox' domain proteins.	BL00027 26.43 1.600e-29 85-128
776	BL01031	Heat shock hsp20 proteins family profile.	BL01031C 17.68 7.000e-13 100-125 BL01031B 15.78 4.300e-11 72-93
777	BL00657	Fork head domain proteins.	BL00657B 22.27 4.789e-37 63-106 BL00657A 19.39 1.600e-32 18-60
782	BL00491	Aminopeptidase P and proline dipeptidase proteins.	BL00491C 12.15 8.800e-18 363-378 BL00491D 8.33 2.946e-12 392-406 BL00491B 5.42 5.320e-12 341-354
783	BL00170	Cyclophilin-type peptidyl-prolyl cis-trans isomerase signatur.	BL00170C 18.49 3.571e-32 35-80
784	BL00226	Intermediate filaments proteins.	BL00226D 19.10 6.143e-40 418-465 BL00226B 23.86 5.696e-35 251-299 BL00226C 13.23 2.174e-23 317-348 BL00226A 12.77 3.571e-12 150-165 BL00226B 23.86 1.113e-10 202-250 BL00226B 23.86 5.395e-09 379-427 BL00226B 23.86 9.163e-09 397-445

SEQ ID NO:	Accession No.	Description	Results*
785	PF00624	Flocculin repeat proteins.	PF00624I 9.10 8.875e-10 96-126
786	BL00021	Kringle domain proteins.	BL00021D 24.56 3.942e-22 376-418 BL00021B 13.33 4.214e-14 217-235
790	BL00027	'Homeobox' domain proteins.	BL00027 26.43 7.750e-34 207-250
791	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 7.000e-14 984-997 PD00066 13.92 1.500e-13 898-911 PD00066 13.92 5.000e-13 956-969 PD00066 13.92 4.429e-12 809-822 PD00066 13.92 3.400e-09 928-941
796	BL01228	Hypothetical cof family proteins.	BL01228D 17.44 7.150e-11 232-257
800	PR00449	TRANSFORMING PROTEIN P21 RAS SIGNATURE	PR00449A 13.20 7.577e-10 21-43
801	BL00741	Guanine-nucleotide dissociation stimulators CDC24 family sign.	BL00741B 14.27 5.250e-10 748-771
802	PR00918	CALICIVIRUS NON-STRUCTURAL POLYPROTEIN FAMILY SIGNATURE	PR00918A 13.76 2.500e-11 1636-1657
811	BL01185	C-terminal cystine knot proteins.	BL01185D 23.45 8.043e-19 4238-4291 BL01185C 15.86 9.852e-15 3615-3654
813	BL00660	Band 4.1 family domain proteins.	BL00660C 23.36 4.774e-17 217-261 BL00660A 31.50 2.091e-16 45-98 BL00660B 17.33 1.396e-09 131-171
814	DM00179	w KINASE ALPHA ADHESION T-CELL.	DM00179 13.97 8.435e-09 17-27
815	BL01205	Iodothyronine deiodinases proteins.	BL01205A 28.90 1.581e-25 12-44
816	BL00126	3'5'-cyclic nucleotide phosphodiesterases proteins.	BL00126C 22.07 1.000e-28 245-286 BL00126E 35.22 6.878e-22 372-427 BL00126D 25.50 1.857e-18 300-339 BL00126A 27.56 4.545e-18 179-216 BL00126B 15.20 2.385e-14 219-231
818	BL01208	VWFC domain proteins.	BL01208B 15.83 5.667e-11 51-66 BL01208B 15.83 7.750e-10 270-285
821	BL01107	Ribosomal protein L27e proteins.	BL01107B 16.28 1.000e-40 46-90 BL01107A 12.03 7.529e-34 3-46
824	PR00621	HISTONE H2B SIGNATURE	PR00621A 12.25 8.714e-23 38-57 PR00621B 4.91 7.207e-21 57-78
827	PR00211	GLUTELIN SIGNATURE	PR00211B 0.86 8.083e-09 102-123
829	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 9.924e-11 19-34
830	BL00226	Intermediate filaments proteins.	BL00226B 23.86 4.600e-33 244-292 BL00226D 19.10 8.054e-29 410-457 BL00226C 13.23 8.125e-22 309-340 BL00226A 12.77 4.960e-14 139-154
833	PR00021	SMALL PROLINE-RICH PROTEIN SIGNATURE	PR00021A 4.31 2.440e-10 2-15 PR00021B 7.29 3.647e-09 24-34
834	PR00876	NEMATODE METALLOTHIONEIN	PR00876B 7.66 5.014e-09 143-157

SEQ ID NO:	Accession No.	Description	Results*
		SIGNATURE	
835	PR00021	SMALL PROLINE-RICH PROTEIN SIGNATURE	PR00021A 4.31 7.366e-16 19-32 PR00021A 4.31 8.291e-09 3-16
836	PF01062	Putative membrane protein.	PF01062F 17.08 1.000e-40 277-331 PF01062E 16.81 8.603e-26 214-258 PF01062D 18.73 8.636e-26 123-167 PF01062A 16.52 6.339e-22 20-60 PF01062B 15.58 6.906e-18 62-92 PF01062C 15.18 5.135e-12 92-123
841	BL00232	Cadherins extracellular repeat proteins domain proteins.	BL00232B 32.79 5.579e-22 18-66 BL00232B 32.79 9.169e-18 236-284 BL00232B 32.79 6.803e-14 340-388 BL00232C 10.65 8.500e-13 234-252 BL00232B 32.79 2.098e-12 120-168 BL00232C 10.65 3.415e-12 16-34 BL00232B 32.79 9.451e-12 451-499
842	PR00021	SMALL PROLINE-RICH PROTEIN SIGNATURE	PR00021A 4.31 5.333e-15 4-17
843	DM00031	IMMUNOGLOBULIN V REGION.	DM00031B 15.41 6.108e-10 91-125
844	BL01242	Formamidopyrimidine-DNA glycosylase proteins.	BL01242F 17.92 7.722e-14 177-211 BL01242G 25.36 3.084e-10 237-281
846	BL00903	Cytidine and deoxycytidylate deaminases zinc-binding region s.	BL00903 12.93 5.821e-09 91-101
848	BL00564	Argininosuccinate synthase proteins.	BL00564A 19.93 6.114e-09 7-44
849	BL00273	Heat-stable enterotoxins proteins.	BL00273 12.24 7.638e-10 140-153 BL00273 12.24 8.875e-10 47-60
850	BL00657	Fork head domain proteins.	BL00657A 19.39 9.438e-21 74-116
855	BL00021	Kringle domain proteins.	BL00021B 13.33 3.143e-18 586-604 BL00021D 24.56 3.613e-17 749-791
860	BL00798	Aldo/keto reductase family proteins.	BL00798F 23.30 1.000e-40 238-287 BL00798E 20.32 8.759e-31 177-215 BL00798B 16.01 3.172e-22 36-61 BL00798D 7.65 1.375e-15 94-111 BL00798A 14.97 2.565e-15 8-23 BL00798C 11.15 2.800e-15 70-83
861	DM01117	2 kw TRANSPOSASE WITHIN TRANSPOSITION VASOTOCIN.	DM01117B 13.11 8.333e-09 495-530
862	PR00930	HIGH MOBILITY GROUP PROTEIN (HMGY) SIGNATURE	PR00930E 5.98 6.143e-09 49-62
863	PD00919	CALCIUM-BINDING PRECURSOR SIGNAL R.	PD00919B 9.47 4.822e-09 171-186
864	BL00021	Kringle domain proteins.	BL00021D 24.56 3.647e-33 490-532
865	PR00910	LUTEOVIRUS ORF6 PROTEIN SIGNATURE	PR00910A 2.51 4.889e-10 56-69
868	PR00833	POLLEN ALLERGEN POA PI SIGNATURE	PR00833H 2.30 8.500e-10 282-297 PR00833H 2.30 6.769e-09 325-340

SEQ ID NO:	Accession No.	Description	Results*
869	BL00032	'Homeobox' antennapedia-type protein.	BL00032B 10.83 1.281e-11 99-138
876	BL00027	'Homeobox' domain proteins.	BL00027 26.43 3.500e-25 177-220
877	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 9.557e-13 134-149 PR00049D 0.00 2.500e-12 136-151 PR00049D 0.00 2.500e-12 137-152 PR00049D 0.00 4.000e-12 138-153 PR00049D 0.00 4.000e-12 139-154 PR00049D 0.00 4.000e-12 140-155 PR00049D 0.00 4.000e-12 141-156 PR00049D 0.00 4.000e-12 142-157 PR00049D 0.00 4.000e-12 143-158 PR00049D 0.00 4.000e-12 144-159 PR00049D 0.00 4.000e-12 145-160 PR00049D 0.00 4.000e-12 146-161 PR00049D 0.00 4.000e-12 147-162 PR00049D 0.00 4.000e-12 148-163 PR00049D 0.00 7.126e-11 132-147 PR00049D 0.00 9.244e-11 149-164 PR00049D 0.00 1.643e-10 135-150 PR00049D 0.00 7.643e-10 131-146 PR00049D 0.00 8.714e-10 133-148 PR00049D 0.00 2.831e-09 130-145 PR00049D 0.00 5.576e-09 150-165
880	PF00624	Flocculin repeat proteins.	PF00624I 9.10 9.646e-09 409-439
881	PF00624	Flocculin repeat proteins.	PF00624I 9.10 9.646e-09 448-478
882	PR00320	G-PROTEIN BETA WD-40 REPEAT SIGNATURE	PR00320B 12.19 3.571e-10 1121-1136 PR00320A 16.74 9.206e-10 1171-1186 PR00320C 13.01 1.000e-09 1121-1136 PR00320A 16.74 1.878e-09 1121-1136 PR00320C 13.01 3.700e-09 1171-1186 PR00320B 12.19 5.950e-09 1171-1186
883	BL00904	Protein prenyltransferases alpha subunit repeat proteins proteins.	BL00904D 1.47 6.945e-10 197-238
884	BL00904	Protein prenyltransferases alpha subunit repeat proteins proteins.	BL00904D 1.47 6.945e-10 179-220
887	PR00254	NICOTINIC ACETYLCHOLINE RECEPTOR SIGNATURE	PR00254D 15.50 1.857e-18 97-116 PR00254A 11.23 2.588e-14 27-44 PR00254C 11.36 3.045e-13 79-92 PR00254B 12.97 5.179e-13 61-76
889	PD02870	RECEPTOR INTERLEUKIN-1 PRECURSOR.	PD02870B 18.83 7.571e-19 101-134 PD02870C 24.41 4.643e-10 146-181
890	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 9.500e-25 148-188 BL00237D 11.23 5.235e-15 381-398 BL00237C 13.19 1.360e-14 319-346 BL00237B 5.28 8.875e-11 276-288

SEQ ID NO:	Accession No.	Description	Results*
892	PD01719	PRECURSOR GLYCOPROTEIN SIGNAL RE.	PD01719A 12.89 8.132e-18 59-87

* Results include: Accession number, sub type, Ematrix p-value, and the position of signature sequence.

TABLE 4

SEQ ID NO:	Pfam Model	Description	E-value	Score
451	tsp_1	Thrombospondin type 1 domain	4.9e-13	56.8
452	proteasome	Proteasome A-type and B-type	2.1e-49	177.6
453	proteasome	Proteasome A-type and B-type	1.5e-39	144.8
454	ldh	lactate/malate dehydrogenase, NAD binding do	1.4e-20	80.1
536	Collagen	Collagen triple helix repeat (20 copies)	1.4e-71	251.2
540	PH	PH domain	7.9e-14	54.2
542	adh_short	short chain dehydrogenase	1.1e-70	248.3
545	UPF0066	Uncharacterised protein family UPF0066	8.1e-38	139.1
546	Peptidase_M48	Peptidase family M48	0.013	-49.3
550	tRNA-synt_1d	tRNA synthetases class I (R)	1.3e-11	17.9
551	mito_carr	Mitochondrial carrier protein	1.3e-20	81.9
554	zf-CCCH	Zinc finger C-x8-C-x5-C-x3-H type	1.2e-09	45.5
559	bZIP	bZIP transcription factor	6e-05	22.9
564	cadherin	Cadherin domain	0	1932.1
565	TGS	TGS domain	0.071	5.1
568	fer2	2Fe-2S iron-sulfur cluster binding domain	2.3e-06	34.6
570	tRNA-synt_2	tRNA synthetases class II (D, K and N)	3.6e-33	123.6
571	SIR2	Sir2 family	1e-97	338.1
574	S-AdoMet_syntD2	S-adenosylmethionine synthetase, cent	1.5e-98	340.9
576	OTU	OTU-like cysteine protease	0.006	13.2
579	R3H	R3H domain	5.5e-14	59.9
580	Ribosomal_L15	Ribosomal protein L15 amino terminal re	4.3e-13	56.9
581	WD40	WD domain, G-beta repeat	1.2e-20	82.1
583	rrm	RNA recognition motif.	5.3e-05	30.1
584	synaptobrevin	Synaptobrevin	5e-36	133.1
585	Oxysterol_BP	Oxysterol-binding protein	7.5e-34	125.9
589	efhand	EF hand	3.4e-26	100.5
590	DPPIV_N_term	Dipeptidyl peptidase IV (DPP IV) N-termi	3.5e-173	588.7
592	WH1	WH1 domain	0.0045	7.1
593	PP2C	Protein phosphatase 2C	1.3e-74	261.3
595	WD40	WD domain, G-beta repeat	3.1e-16	67.4
596	ras	Ras family	2.6e-86	300.1
597	filament	Intermediate filament protein	1.5e-06	28.4
598	zf-C2H2	Zinc finger, C2H2 type	1.4e-106	367.4
599	thiored	Thioredoxin	8.9e-46	156.1
603	PAP2	PAP2 superfamily	0.0057	10.0
604	PDZ	PDZ domain (Also known as DHR or GLGF)	3.4e-23	90.5
606	NAC	NAC domain	1.6e-26	101.5
607	Anti_proliferat	BTG1 family	5.2e-22	86.5
609	Ribosomal_S27e	Ribosomal protein S27	8.4e-30	112.4
611	DSPc	Dual specificity phosphatase, catalytic doma	2.4e-06	23.2
612	ig	Immunoglobulin domain	9e-10	36.5
613	Gal-bind_lectin	Galactoside-binding lectin	1.9e-07	20.0
614	Collagen	Collagen triple helix repeat (20 copies)	9e-41	148.9
616	PHD	PHD-finger	4.9e-20	80.0
618	bZIP	bZIP transcription factor	0.0062	15.8
620	AP_endonucleas1	AP endonuclease family 1	0.021	10.3
622	SET	SET domain	2e-54	194.2
623	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	1.3e-10	38.7
624	ACBP	Acyl CoA binding protein	4.4e-57	203.1

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE 5

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
449	1avl	A	188	388	5.4e-07			77.81	APOLIPOPROTEIN A-I; CHAIN: A, B, C, D;	LIPID TRANSPORT APO A-I; LIPOPROTEIN, LIPID TRANSPORT, CHOLESTEROL METABOLISM, 2 ATHEROSCLEROSIS, HDL, LCAT-ACTIVATION
449	1cun	A	133	349	3.6e-15	0.10	0.19		ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
449	1cun	A	154	364	3.6e-15			72.53	ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
449	1dn1	B	149	374	3.6e-17	-0.44	0.10		SYNTAXIN BINDING PROTEIN 1; CHAIN: A; SYNTAXIN 1A; CHAIN: B;	ENDOCYTOSIS/EXOCYTOSIS NSEC1; PROTEIN-PROTEIN COMPLEX, MULTI-SUBUNIT
449	1dn1	B	227	412	1.8e-15	-0.04	0.35		SYNTAXIN BINDING PROTEIN 1; CHAIN: A; SYNTAXIN 1A; CHAIN: B;	ENDOCYTOSIS/EXOCYTOSIS NSEC1; PROTEIN-PROTEIN COMPLEX, MULTI-SUBUNIT
449	1ez3	A	142	291	9e-09	0.04	-0.08		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
449	1ez3	A	149	273	1.4e-09	0.03	-0.06		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
449	1ez3	A	177	301	3.6e-09	0.14	-0.08		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
449	1ez3	A	192	338	3.6e-09	-0.09	0.12		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
449	1ez3	A	240	374	1.3e-10	0.06	-0.06		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
449	1ez3	A	263	380	9e-11	0.12	-0.14		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
449	1qqe	A	112	388	1.1e-15			71.23	VESICULAR TRANSPORT PROTEIN SEC17; CHAIN: A;	PROTEIN TRANSPORT HELIX-TURN-HELIX TPR-LIKE REPEAT, PROTEIN TRANSPORT
449	1quu	A	144	408	1.8e-24	-0.03	0.35		HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
449	1quu	A	154	407	1.8e-24			74.70	HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
449	1sig		118	308	7.2e-07	-0.31	0.16		RNA POLYMERASE PRIMARY SIGMA FACTOR; CHAIN: NULL;	TRANSCRIPTION REGULATION SIGMA70; RNA POLYMERASE SIGMA FACTOR, TRANSCRIPTION REGULATION
449	1sig		141	440	3.6e-12			78.55	RNA POLYMERASE PRIMARY SIGMA FACTOR; CHAIN: NULL;	TRANSCRIPTION REGULATION SIGMA70; RNA POLYMERASE SIGMA FACTOR, TRANSCRIPTION REGULATION
449	1av1	A	159	357	5.4e-11			72.01	APOLIPOPROTEIN A-I; CHAIN: A, B, C, D;	LIPID TRANSPORT APO A-I; LIPOPROTEIN, LIPID TRANSPORT, CHOLESTEROL METABOLISM, 2 ATHEROSCLEROSIS, HDL, LCAT-ACTIVATION
449	1cun	A	122	307	7.2e-13	0.06	0.40		ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
449	1cun	A	133	360	1.6e-15	0.10	0.84		ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
449	1cun	A	154	370	1.6e-15			64.60	ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
449	1cun	A	82	271	9e-10	-0.38	0.01		ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
449	1dn1	B	173	381	3.6e-14	-0.24	0.05		SYNTAXIN BINDING PROTEIN 1; CHAIN: A; SYNTAXIN 1A; CHAIN: B;	ENDOCYTOSIS/EXOCYTOSIS NSEC1; PROTEIN-PROTEIN COMPLEX, MULTI-SUBUNIT
449	1e94	E	99	294	1.8e-05	-0.36	0.23		HEAT SHOCK PROTEIN HSLV;	CHAPERONE HSLV; HSLU CHAPERONE,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: A, B, C, D; HEAT SHOCK PROTEIN HSLU; CHAIN: E, F;	HSLVU, CLPQY, AAA-ATPASE, ATP-DEPENDENT 2 PROTEOLYSIS, PROTEASOME
449	1ez3	A	226	349	1.8e-09	0.07	-0.13		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
449	1ez3	A	246	374	5.4e-11	0.06	-0.13		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
449	1qqe	A	112	400	3.6e-13			67.33	VESICULAR TRANSPORT PROTEIN SEC17; CHAIN: A;	PROTEIN TRANSPORT HELIX-TURN-HELIX TPR-LIKE REPEAT, PROTEIN TRANSPORT
449	1quu	A	134	377	7.2e-23	-0.13	0.43		HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
449	1quu	A	154	399	7.2e-23			66.85	HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
449	1sig		67	372	1.6e-10			71.14	RNA POLYMERASE PRIMARY SIGMA FACTOR; CHAIN: NULL;	TRANSCRIPTION REGULATION SIGMA70; RNA POLYMERASE SIGMA FACTOR, TRANSCRIPTION REGULATION
450	1av1	A	188	388	5.4e-07			77.81	APOLIPOPROTEIN A-I; CHAIN: A, B, C, D;	LIPID TRANSPORT APO A-I; LIPOPROTEIN, LIPID TRANSPORT, CHOLESTEROL METABOLISM, 2 ATHEROSCLEROSIS, HDL, LCAT-ACTIVATION
450	1cun	A	133	349	3.6e-15	0.10	0.19		ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
450	1cun	A	154	364	3.6e-15			72.53	ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
450	1dn1	B	149	374	3.6e-17	-0.44	0.10		SYNTAXIN BINDING PROTEIN 1; CHAIN: A; SYNTAXIN 1A; CHAIN: B;	ENDOCYTOSIS/EXOCYTOSIS NSEC1; PROTEIN-PROTEIN COMPLEX, MULTI-SUBUNIT
450	1dn1	B	227	412	1.8e-15	-0.04	0.35		SYNTAXIN BINDING PROTEIN 1; CHAIN: A; SYNTAXIN 1A; CHAIN:	ENDOCYTOSIS/EXOCYTOSIS NSEC1; PROTEIN-PROTEIN COMPLEX, MULTI-

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									B;	SUBUNIT
450	1ez3	A	142	291	9e-09	0.04	-0.08		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
450	1ez3	A	149	273	1.4e-09	0.03	-0.06		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
450	1ez3	A	177	301	3.6e-09	0.14	-0.08		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
450	1ez3	A	192	338	3.6e-09	-0.09	0.12		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
450	1ez3	A	240	374	1.3e-10	0.06	-0.06		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
450	1ez3	A	263	380	9e-11	0.12	-0.14		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
450	1qqe	A	112	388	1.1e-15			71.23	VESICULAR TRANSPORT PROTEIN SEC17; CHAIN: A;	PROTEIN TRANSPORT HELIX-TURN-HELIX TPR-LIKE REPEAT, PROTEIN TRANSPORT
450	1quu	A	144	408	1.8e-24	-0.03	0.35		HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
450	1quu	A	154	407	1.8e-24			74.70	HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
450	1sig		118	308	7.2e-07	-0.31	0.16		RNA POLYMERASE PRIMARY SIGMA FACTOR; CHAIN: NULL;	TRANSCRIPTION REGULATION SIGMA70; RNA POLYMERASE SIGMA FACTOR, TRANSCRIPTION REGULATION
450	1sig		141	440	3.6e-12			78.55	RNA POLYMERASE PRIMARY SIGMA FACTOR; CHAIN: NULL;	TRANSCRIPTION REGULATION SIGMA70; RNA POLYMERASE SIGMA FACTOR, TRANSCRIPTION REGULATION
450	1av1	A	159	357	5.4e-11			72.01	APOLIPOPROTEIN A-I; CHAIN: A, B, C, D;	LIPID TRANSPORT APO A-I; LIPOPROTEIN, LIPID TRANSPORT, CHOLESTEROL METABOLISM, 2 ATHEROSCLEROSIS, HDL, LCAT-ACTIVATION
450	1cun	A	122	307	7.2e-13	0.06	0.40		ALPHA SPECTRIN; CHAIN: A, B,	STRUCTURAL PROTEIN TWO REPEATS OF

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									C;	SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
450	1cun	A	133	360	1.6e-15	0.10	0.84		ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
450	1cun	A	154	370	1.6e-15			64.60	ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
450	1cun	A	82	271	9e-10	-0.38	0.01		ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
450	1dn1	B	173	381	3.6e-14	-0.24	0.05		SYNTAXIN BINDING PROTEIN 1; CHAIN: A; SYNTAXIN 1A; CHAIN: B;	ENDOCYTOSIS/EXOCYTOSIS NSEC1; PROTEIN-PROTEIN COMPLEX, MULTI-SUBUNIT
450	1e94	E	99	294	1.8e-05	-0.36	0.23		HEAT SHOCK PROTEIN HSLV; CHAIN: A, B, C, D; HEAT SHOCK PROTEIN HSLU; CHAIN: E, F;	CHAPERONE HSLV; HSLU CHAPERONE, HSLVU, CLPQY, AAA-ATPASE, ATP-DEPENDENT 2 PROTEOLYSIS, PROTEASOME
450	1ez3	A	226	349	1.8e-09	0.07	-0.13		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
450	1ez3	A	246	374	5.4e-11	0.06	-0.13		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
450	1qqe	A	112	400	3.6e-13			67.33	VESICULAR TRANSPORT PROTEIN SEC17; CHAIN: A;	PROTEIN TRANSPORT HELIX-TURN-HELIX TPR-LIKE REPEAT, PROTEIN TRANSPORT
450	1quu	A	134	377	7.2e-23	-0.13	0.43		HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
450	1quu	A	154	399	7.2e-23			66.85	HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
450	1sig		67	372	1.6e-10			71.14	RNA POLYMERASE PRIMARY SIGMA FACTOR; CHAIN: NULL;	TRANSCRIPTION REGULATION SIGMA70; RNA POLYMERASE SIGMA FACTOR, TRANSCRIPTION REGULATION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
451	9wga	A	688	842	1.6e-15	0.02	-0.19		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
452	1ryp	C	2	237	8e-73	0.76	1.00		20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
452	1ryp	C	2	240	1.6e-75	0.83	1.00		20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
452	1ryp	C	2	243	1.6e-75			245.13	20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
452	1g0u	D	16	193	1.8e-43	0.39	1.00		PROTEASOME COMPONENT Y7; CHAIN: A, O; PROTEASOME COMPONENT Y13; CHAIN: B, P; PROTEASOME COMPONENT PRE6; CHAIN: C, Q; PROTEASOME COMPONENT PUP2; CHAIN: D, R; PROTEASOME COMPONENT PRE5; CHAIN: E, S; PROTEASOME COMPONENT C1; CHAIN: F, T; PROTEASOME COMPONENT C7-ALPHA; CHAIN: G, U; PROTEASOME COMPONENT PUP1; CHAIN: H, V; PROTEASOME COMPONENT PUP3; CHAIN: I, W; PROTEASOME COMPONENT C11; CHAIN: J, X; PROTEASOME COMPONENT PRE2; CHAIN: K, Y; PROTEASOME COMPONENT C5;	HYDROLASE MACROPAIN SUBUNIT Y7, PROTEINASE YSCE SUBUNIT 7, MACROPAIN SUBUNIT Y13, PROTEINASE YSCE SUBUNIT 13, MACROPAIN SUBUNIT PRE6, PROTEINASE YSCE SUBUNIT MACROPAIN SUBUNIT PUP2, PROTEINASE YSCE SUBUNIT MACROPAIN SUBUNIT PRE5, PROTEINASE YSCE SUBUNIT MACROPAIN SUBUNIT C1, PROTEINASE YSCE SUBUNIT 1, MACROPAIN SUBUNIT C7-ALPHA, PROTEINASE YSCE MACROPAIN SUBUNIT PUP1, PROTEINASE YSCE SUBUNIT MACROPAIN SUBUNIT PUP3, MULTICATALYTIC MACROPAIN SUBUNIT C11, PROTEINASE YSCE SUBUNIT 11, MACROPAIN SUBUNIT PRE2,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: L, Z; PROTEASOME COMPONENT PRE4; CHAIN: M, 1; PROTEASOME COMPONENT PRE3; CHAIN: N, 2;	PROTEINASE YSCE SUBUNIT MULTICATALYTIC ENDOPEPTIDASE COMPLEX SUBUNIT C5; MACROPAIN SUBUNIT PRE4, PROTEINASE YSCE SUBUNIT MACROPAIN SUBUNIT PRE3, PROTEINASE YSCE SUBUNIT PROTEASOME, UBIQUITIN, DEGRADATION, PROTEASE, NTN-HYDROLASE
452	1pma	A	1	206	3.2e-44	0.55	1.00		PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	PROTEASE PROSOME, MULTICATALYTIC PROTEASE, MCP, MACROPAIN; PROTEASE, PROTEASOME, HYDROLASE
452	1pma	A	3	206	3.2e-44			105.95	PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	PROTEASE PROSOME, MULTICATALYTIC PROTEASE, MCP, MACROPAIN; PROTEASE, PROTEASOME, HYDROLASE
452	1ryp	B	1	206	1.3e-44	0.67	1.00		20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
452	1ryp	B	1	216	1.3e-44			107.59	20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
452	1ryp	C	2	206	4.8e-51	0.73	1.00		20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
452	1ryp	C	2	209	5.4e-54	0.48	1.00		20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
452	1ryp	C	2	212	5.4e-54			172.82	20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
453	1ryp	C	2	237	8e-73	0.76	1.00		20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
453	1ryp	C	2	240	1.6e-75	0.83	1.00		20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
453	1ryp	C	2	243	1.6e-75			245.13	20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
453	1g0u	D	16	193	1.8e-43	0.39	1.00		PROTEASOME COMPONENT Y7; CHAIN: A, O; PROTEASOME COMPONENT Y13; CHAIN: B, P; PROTEASOME COMPONENT PRE6; CHAIN: C, Q; PROTEASOME COMPONENT PUP2; CHAIN: D, R; PROTEASOME COMPONENT PRE5; CHAIN: E, S; PROTEASOME COMPONENT C1; CHAIN: F, T; PROTEASOME COMPONENT C7-ALPHA; CHAIN: G, U; PROTEASOME COMPONENT PUP1; CHAIN: H, V; PROTEASOME COMPONENT PUP3; CHAIN: I, W; PROTEASOME COMPONENT C11; CHAIN: J, X; PROTEASOME COMPONENT PRE2; CHAIN: K, Y; PROTEASOME COMPONENT C5; CHAIN: L, Z; PROTEASOME COMPONENT PRE4; CHAIN: M, 1;	HYDROLASE MACROPAIN SUBUNIT Y7, PROTEINASE YSCE SUBUNIT 7, MACROPAIN SUBUNIT Y13, PROTEINASE YSCE SUBUNIT 13, MACROPAIN SUBUNIT PRE6, PROTEINASE YSCE SUBUNIT MACROPAIN SUBUNIT PUP2, PROTEINASE YSCE SUBUNIT MACROPAIN SUBUNIT PRE5, PROTEINASE YSCE SUBUNIT MACROPAIN SUBUNIT C1, PROTEINASE YSCE SUBUNIT 1, MACROPAIN SUBUNIT C7-ALPHA, PROTEINASE YSCE MACROPAIN SUBUNIT PUP1, PROTEINASE YSCE SUBUNIT MACROPAIN SUBUNIT PUP3, MULTICATALYTIC MACROPAIN SUBUNIT C11, PROTEINASE YSCE SUBUNIT 11, MACROPAIN SUBUNIT PRE2, PROTEINASE YSCE SUBUNIT MULTICATALYTIC ENDOPEPTIDASE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									PROTEASOME COMPONENT PRE3; CHAIN: N, 2;	COMPLEX SUBUNIT C5; MACROPAIN SUBUNIT PRE4, PROTEINASE YSCE SUBUNIT MACROPAIN SUBUNIT PRE3, PROTEINASE YSCE SUBUNIT PROTEASOME, UBIQUITIN, DEGRADATION, PROTEASE, NTN-HYDROLASE
453	1pma	A	1	206	3.2e-44	0.55	1.00		PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	PROTEASE PROSOME, MULTICATALYTIC PROTEASE, MCP, MACROPAIN; PROTEASE, PROTEASOME, HYDROLASE
453	1pma	A	3	206	3.2e-44			105.95	PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	PROTEASE PROSOME, MULTICATALYTIC PROTEASE, MCP, MACROPAIN; PROTEASE, PROTEASOME, HYDROLASE
453	1ryp	B	1	206	1.3e-44	0.67	1.00		20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
453	1ryp	B	1	216	1.3e-44			107.59	20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
453	1ryp	C	2	206	4.8e-51	0.73	1.00		20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
453	1ryp	C	2	209	5.4e-54	0.48	1.00		20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE
453	1ryp	C	2	212	5.4e-54			172.82	20S PROTEASOME; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	MULTICATALYTIC PROTEINASE MULTICATALYTIC PROTEINASE, 20S PROTEASOME, PROTEIN 2 DEGRADATION, ANTIGEN PROCESSING, HYDROLASE, PROTEASE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
454	1a5z		185	360	6.4e-53	0.23	0.93		L-LACTATE DEHYDROGENASE; CHAIN: NULL;	OXIDOREDUCTASE OXIDOREDUCTASE, GLYCOLYSIS, HYPERTHERMOPHILES, THERMOTOGA 2 MARITIMA, PROTEIN STABILITY
454	1f0y	A	184	347	1.6e-18	-0.35	0.62		L-3-HYDROXYACYL-COA DEHYDROGENASE; CHAIN: A, B;	OXIDOREDUCTASE HCDH; ABORTIVE TERNARY COMPLEX
454	1ldb		177	378	9.6e-48	0.20	1.00		OXIDOREDUCTASE(CHOH(D)-NAD(A)) APO-*L-*LACTATE DEHYDROGENASE (E.C.1.1.1.27) 1LDB 4	
454	1ldn	A	177	368	1.6e-52	0.12	1.00		OXIDOREDUCTASE(CHOH(D)-NAD(A)) L-LACTATE DEHYDROGENASE (E.C.1.1.1.27) COMPLEXED WITH NADH, 1LDN 3 OXAMATE, AND FRUCTOSE-1,6-BISPHOSPHATE 1LDN 4	
454	1llc		175	379	9.6e-52			51.38	OXIDOREDUCTASE(CHOH(D)-NAD(A)) L-*LACTATE DEHYDROGENASE (E.C.1.1.1.27) COMPLEX WITH 1LLC 4 FRUCTOSE-1,6-BISPHOSPHATE (/FBP$) AND CO==2+== 1LLC 5	
454	1llc		182	376	9.6e-52	0.00	0.47		OXIDOREDUCTASE(CHOH(D)-NAD(A)) L-*LACTATE DEHYDROGENASE (E.C.1.1.1.27) COMPLEX WITH 1LLC 4 FRUCTOSE-1,6-BISPHOSPHATE (/FBP$) AND CO==2+== 1LLC 5	
454	1lld	A	189	368	3.2e-47	0.01	0.90		OXIDOREDUCTASE(CHOH (D)-NAD (A)) L-LACTATE DEHYDROGENASE (E.C.1.1.1.27) (T-STATE) MUTANT 1LLD 3 WITH CYS 199 REPLACED BY SER (C199S) COMPLEX WITH NADH 1LLD 4	
454	2aak		16	171	4.8e-44	0.02	-0.13		UBIQUITIN CONJUGATING ENZYME; CHAIN: NULL;	UBIQUITIN CONJUGATION UBC1; UBIQUITIN CONJUGATION, LIGASE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
454	2cmd		182	378	3.2e-47	0.09	0.48		OXIDOREDUCTASE(NAD(A)-CHOH(D)) MALATE DEHYDROGENASE (E.C.1.1.1.37) 2CMD 3	
454	2ldx		164	378	9.6e-59			81.97	OXIDOREDUCTASE(CHOH(D)-NAD(A)) APO-LACTATE DEHYDROGENASE (E.C.1.1.1.27), ISOENZYME C=4=2LDX 4	
454	2ldx		168	374	9.6e-59	0.08	1.00		OXIDOREDUCTASE(CHOH(D)-NAD(A)) APO-LACTATE DEHYDROGENASE (E.C.1.1.1.27), ISOENZYME C=4=2LDX 4	
454	3hdh	A	184	347	8e-18	-0.46	0.46		L-3-HYDROXYACYL COA DEHYDROGENASE; CHAIN: A, B, C;	OXIDOREDUCTASE SCHAD; OXIDOREDUCTASE, BETA OXIDATION, SCHAD, CATALYTIC ACTIVITY: 2 L-3-HYDROXYACYL-COA + NAD(+) = 3-OXOACYL-COA + NADH
454	3hdh	C	184	347	8e-18	0.01	0.35		L-3-HYDROXYACYL COA DEHYDROGENASE; CHAIN: A, B, C;	OXIDOREDUCTASE SCHAD; OXIDOREDUCTASE, BETA OXIDATION, SCHAD, CATALYTIC ACTIVITY: 2 L-3-HYDROXYACYL-COA + NAD(+) = 3-OXOACYL-COA + NADH
454	5ldh		164	376	8e-60			80.62	OXIDOREDUCTASE, CHOH DONOR, NAD ACCEPTR LACTATE DEHYDROGENASE H=4= AND S-$LAC-/NAD$=+= COMPLEX 5LDH 4 (E.C.1.1.1.27) 5LDH 5	
454	5ldh		185	375	8e-60	0.02	1.00		OXIDOREDUCTASE, CHOH DONOR, NAD ACCEPTR LACTATE DEHYDROGENASE H=4= AND S-$LAC-/NAD$=+= COMPLEX 5LDH 4 (E.C.1.1.1.27) 5LDH 5	
454	6ldh		164	378	1.6e-57			74.66	OXIDOREDUCTASE(CHOH(D)-NAD(A)) M=4= APO-*LACTATE DEHYDROGENASE (E.C.1.1.1.27) 6LDH 4	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
454	6ldh		172	378	1.6e-57	0.26	1.00		OXIDOREDUCTASE(CHOH(D)-NAD(A)) M=4= APO-*LACTATE DEHYDROGENASE (E.C.1.1.1.27) 6LDH 4	
454	9ldt	A	164	377	1.4e-61			78.35	OXIDOREDUCTASE(CHOH(D)-NAD+(A)) LACTATE DEHYDROGENASE (E.C.1.1.1.27) COMPLEX WITH NADH 9LDT 3 AND OXAMATE 9LDT 4	
454	9ldt	A	172	375	1.4e-61	0.19	1.00		OXIDOREDUCTASE(CHOH(D)-NAD+(A)) LACTATE DEHYDROGENASE (E.C.1.1.1.27) COMPLEX WITH NADH 9LDT 3 AND OXAMATE 9LDT 4	
458	1avl	A	21	219	5.4e-05			54.33	APOLIPOPROTEIN A-I; CHAIN: A, B, C, D;	LIPID TRANSPORT APO A-I; LIPOPROTEIN, LIPID TRANSPORT, CHOLESTEROL METABOLISM, 2 ATHEROSCLEROSIS, HDL, LCAT-ACTIVATION
458	1cun	A	25	222	1.8e-07			58.19	ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
458	1dn1	B	20	186	9e-09	-0.00	-0.12		SYNTAXIN BINDING PROTEIN 1; CHAIN: A; SYNTAXIN 1A; CHAIN: B;	ENDOCYTOSIS/EXOCYTOSIS NSEC1; PROTEIN-PROTEIN COMPLEX, MULTI-SUBUNIT
458	1quu	A	16	222	5.4e-09			56.88	HUMAN SKELETAL MUSCLE ALPHA-ACTININ 2; CHAIN: A;	CONTRACTILE PROTEIN TRIPLE-HELIX COILED COIL, CONTRACTILE PROTEIN
459	1klo		87	243	4.8e-15	0.05	-0.18		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
459	1xka	L	171	247	1.6e-12	0.09	0.00		BLOOD COAGULATION FACTOR XA; CHAIN: L, C;	BLOOD COAGULATION FACTOR STUART FACTOR; BLOOD COAGULATION FACTOR, SERINE PROTEINASE, EPIDERMAL 2 GROWTH FACTOR LIKE DOMAIN
460	1a06		2	318	1.4e-89			114.93	CALCIUM/CALMODULIN-	KINASE KINASE, SIGNAL

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									DEPENDENT PROTEIN KINASE; CHAIN: NULL;	TRANSDUCTION, CALCIUM/CALMODULIN
460	1a06		4	310	1.4e-89	0.37	1.00		CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE; CHAIN: NULL;	KINASE KINASE, SIGNAL TRANSDUCTION, CALCIUM/CALMODULIN
460	1apm	E	1	330	0	0.63	1.00		TRANSFERASE(PHOSPHOTRANS FERASE) $C-/AMP$-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (C/APK) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (/S139A$) COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	
460	1apm	E	1	334	0			110.33	TRANSFERASE(PHOSPHOTRANS FERASE) $C-/AMP$-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (C/APK) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (/S139A$) COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	
460	1aq1		1	291	1.1e-57	0.46	1.00		CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE, PHOSPHORYLATION, STAUROSPORINE, 2 CELL DIVISION, MITOSIS, INHIBITION
460	1aq1		1	292	1.1e-57			102.26	CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE, PHOSPHORYLATION, STAUROSPORINE, 2 CELL DIVISION, MITOSIS, INHIBITION
460	1bi8	A	2	304	8e-47			88.87	CYCLIN-DEPENDENT KINASE 6; CHAIN: A, C; CYCLIN-DEPENDENT KINASE INHIBITOR;	COMPLEX (KINASE/INHIBITOR) CDK6; P19INK4D; CYCLIN DEPENDENT KINASE, CYCLIN DEPENDENT KINASE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: B, D;	INHIBITORY 2 PROTEIN, CDK, INK4, CELL CYCLE, COMPLEX (KINASE/INHIBITOR) HEADER HELIX
460	1bi8	A	4	281	8e-47	0.23	0.99		CYCLIN-DEPENDENT KINASE 6; CHAIN: A, C; CYCLIN-DEPENDENT KINASE INHIBITOR; CHAIN: B, D;	COMPLEX (KINASE/INHIBITOR) CDK6; P19INK4D; CYCLIN DEPENDENT KINASE, CYCLIN DEPENDENT KINASE INHIBITORY 2 PROTEIN, CDK, INK4, CELL CYCLE, COMPLEX (KINASE/INHIBITOR) HEADER HELIX
460	1blx	A	2	308	8e-49			92.84	CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
460	1blx	A	4	282	8e-49	0.44	1.00		CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
460	1byg	A	1	285	3.2e-33			77.17	C-TERMINAL SRC KINASE; CHAIN: A;	TRANSFERASE CSK; PROTEIN KINASE, C-TERMINAL SRC KINASE, PHOSPHORYLATION, 2 STAUROSPORINE, TRANSFERASE
460	1cki	A	2	300	3.6e-45			79.57	CASEIN KINASE I DELTA; 1CKI 6 CHAIN: A, B; 1CKI 7	PHOSPHOTRANSFERASE PROTEIN KINASE 1CKI 18
460	1cki	A	4	285	3.6e-45	0.30	1.00		CASEIN KINASE I DELTA; 1CKI 6 CHAIN: A, B; 1CKI 7	PHOSPHOTRANSFERASE PROTEIN KINASE 1CKI 18
460	1cm8	A	18	280	6.4e-44	0.45	0.95		PHOSPHORYLATED MAP KINASE P38-GAMMA; CHAIN: A, B;	TRANSFERASE STRESS-ACTIVATED PROTEIN KINASE-3, ERK6, ERK5; P38-GAMMA, GAMMA, PHOSPHORYLATION, MAP KINASE
460	1cmk	E	1	330	0	0.42	1.00		PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	
460	1cmk	E	1	334	0			103.35	PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									(E.C.2.7.1.37) 1CMK 4	
460	1ctp	E	1	316	0			100.65	TRANSFERASE(PHOSPHOTRANS FERASE) CAMP-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (CAPK) 1CTP 3 (CATALYTIC SUBUNIT) 1CTP 4	
460	1ctp	E	1	325	0	0.43	1.00		TRANSFERASE(PHOSPHOTRANS FERASE) CAMP-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (CAPK) 1CTP 3 (CATALYTIC SUBUNIT) 1CTP 4	
460	1f3m	C	2	281	3.2e-61	0.68	1.00		SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: A, B; SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: C, D;	TRANSFERASE KINASE DOMAIN, AUTOINHIBITORY FRAGMENT, HOMODIMER
460	1fgk	A	2	286	1.6e-31			86.13	FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
460	1fgk	B	1	285	1.1e-36			80.44	FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
460	1hcl		1	291	4.8e-60	0.40	1.00		HUMAN CYCLIN-DEPENDENT KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; TRANSFERASE, SERINE/THREONINE PROTEIN KINASE, ATP-BINDING, 2 CELL CYCLE, CELL DIVISION, MITOSIS, PHOSPHORYLATION
460	1hcl		1	292	4.8e-60			116.60	HUMAN CYCLIN-DEPENDENT KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; TRANSFERASE, SERINE/THREONINE PROTEIN KINASE, ATP-BINDING, 2 CELL CYCLE, CELL DIVISION, MITOSIS, PHOSPHORYLATION
460	1jnk		1	296	3.2e-46	0.09	0.30		C-JUN N-TERMINAL KINASE; CHAIN: NULL;	TRANSFERASE JNK3; TRANSFERASE, JNK3 MAP KINASE, SERINE/THREONINE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										PROTEIN 2 KINASE
460	1jnk		1	316	3.2e-46			86.64	C-JUN N-TERMINAL KINASE; CHAIN: NULL;	TRANSFERASE JNK3; TRANSFERASE, JNK3 MAP KINASE, SERINE/THREONINE PROTEIN 2 KINASE
460	1koa		1	307	4.8e-70	0.50	1.00		TWITCHIN; CHAIN: NULL;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
460	1kob	A	1	342	6.4e-71			112.15	TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
460	1kob	A	2	311	6.4e-71	0.36	1.00		TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
460	1p38		2	306	9.6e-50	0.15	0.99		MAP KINASE P38; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE; TRANSFERASE, MAP KINASE, SERINE/THREONINE-PROTEIN KINASE, 2 P38
460	1p38		2	349	9.6e-50			81.51	MAP KINASE P38; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE; TRANSFERASE, MAP KINASE, SERINE/THREONINE-PROTEIN KINASE, 2 P38
460	1phk		1	284	1.6e-84			110.57	PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE-PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
460	1phk		2	281	1.6e-84	0.70	1.00		PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE-PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
460	1pme		15	302	9.6e-46	0.45	1.00		ERK2; CHAIN: NULL;	TRANSFERASE MAP KINASE, SERINE/THREONINE PROTEIN KINASE, TRANSFERASE
460	1pme		2	327	9.6e-46			99.96	ERK2; CHAIN: NULL;	TRANSFERASE MAP KINASE, SERINE/THREONINE PROTEIN KINASE, TRANSFERASE
460	1tki	A	1	344	4.8e-58			113.99	TITIN; CHAIN: A, B;	SERINE KINASE SERINE KINASE, TITIN, MUSCLE, AUTOINHIBITION
460	1tki	A	2	281	4.8e-58	0.58	1.00		TITIN; CHAIN: A, B;	SERINE KINASE SERINE KINASE, TITIN, MUSCLE, AUTOINHIBITION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
460	3erk		2	325	6.4e-49			101.77	EXTRACELLULAR REGULATED KINASE 2; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE, MAP 2, ERK2; TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, MAP KINASE, 2 ERK2
460	3erk		3	314	6.4e-49	0.44	1.00		EXTRACELLULAR REGULATED KINASE 2; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE, MAP 2, ERK2; TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, MAP KINASE, 2 ERK2
462	1efn	A	276	328	8e-19	-0.14	1.00		FYN TYROSINE KINASE; CHAIN: A, C; HIV-1 NEF PROTEIN; CHAIN: B, D;	COMPLEX (SH3 DOMAIN/VIRAL ENHANCER) SRC-HOMOLOGY 3 DOMAIN; COMPLEX (SH3 DOMAIN/VIRAL ENHANCER), PROTO-ONCOGENE, 2 TRANSFERASE, TYROSINE-PROTEIN KINASE, PHOSPHORYLATION, 3 AIDS, MYRISTYLATION, GTP-BINDING, ATP-BINDING, SH3 DOMAIN, 4 SH2 DOMAIN, PPII HELIX, PXXP MOTIF
462	1fmk		2	184	9.6e-45	0.22	1.00		TYROSINE-PROTEIN KINASE SRC; CHAIN: NULL;	PHOSPHOTRANSFERASE C-SRC, P60-SRC; SRC, TYROSINE KINASE, PHOSPHORYLATION, SH2, SH3, 2 PHOSPHOTYROSINE, PROTO-ONCOGENE, PHOSPHOTRANSFERASE
462	1fmk		273	327	4.8e-17	-0.02	1.00		TYROSINE-PROTEIN KINASE SRC; CHAIN: NULL;	PHOSPHOTRANSFERASE C-SRC, P60-SRC; SRC, TYROSINE KINASE, PHOSPHORYLATION, SH2, SH3, 2 PHOSPHOTYROSINE, PROTO-ONCOGENE, PHOSPHOTRANSFERASE
462	1fyn	A	273	328	1.4e-19	-0.08	1.00		PHOSPHOTRANSFERASE FYN; CHAIN: A; 3BP-2; CHAIN: B;	TRANSFERASE PROTO-ONCOGENE TYROSINE KINASE; PROTO-ONCOGENE, TRANSFERASE, TYROSINE-PROTEIN KINASE, 2 PHOSPHORYLATION, ATP-BINDING, MYRISTYLATION, SH3 DOMAIN, 3 COMPLEX (PHOSPHOTRANSFERASE/PEPTIDE)
462	1gbr	A	268	327	1.8e-17	-0.03	0.99		SIGNAL TRANSDUCTION PROTEIN GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									(GRB2, N-TERMINAL 1GBR 3 SH3 DOMAIN) COMPLEXED WITH SOS-A PEPTIDE 1GBR 4 (NMR, 29 STRUCTURES) 1GBR 5	
462	1gfc		272	330	4.8e-21			63.10	ADAPTOR PROTEIN CONTAINING SH2 AND SH3 GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2 (GRB2) 1GFC 3 (C-TERMINAL SH3 DOMAIN) (NMR, MINIMIZED MEAN STRUCTURE) 1GFC 4	
462	1gfc		275	330	4.8e-21	-0.00	1.00		ADAPTOR PROTEIN CONTAINING SH2 AND SH3 GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2 (GRB2) 1GFC 3 (C-TERMINAL SH3 DOMAIN) (NMR, MINIMIZED MEAN STRUCTURE) 1GFC 4	
462	1ghu		56	149	5.4e-27			98.07	GRB2; CHAIN: NULL;	SRC HOMOLOGY 2 DOMAIN GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2; SRC HOMOLOGY 2 DOMAIN, GRB2, SH2
462	1ghu		58	148	5.4e-27	0.92	1.00		GRB2; CHAIN: NULL;	SRC HOMOLOGY 2 DOMAIN GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2; SRC HOMOLOGY 2 DOMAIN, GRB2, SH2
462	1ghu		58	148	9.6e-12	1.21	1.00		GRB2; CHAIN: NULL;	SRC HOMOLOGY 2 DOMAIN GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2; SRC HOMOLOGY 2 DOMAIN, GRB2, SH2
462	1gri	A	1	155	1.3e-28	0.81	1.00		GROWTH FACTOR BOUND PROTEIN 2; 1GRI 5 CHAIN: A, B; 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
462	1gri	A	1	216	1.3e-28			161.93	GROWTH FACTOR BOUND PROTEIN 2; 1GRI 5 CHAIN: A, B; 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
462	1gri	A	270	330	8e-22	0.33	1.00		GROWTH FACTOR BOUND PROTEIN 2; 1GRI 5 CHAIN: A, B; 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
462	1lck	A	2	149	9e-31			88.47	P56==LCK== TYROSINE KINASE; 1LCK 7 CHAIN: A; 1LCK 8 TAIL	COMPLEX (KINASE/PEPTIDE)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									PHOSPHOPEPTIDE TEGQ(PHOSPHO)YQPQPA; 1LCK 14 CHAIN: B; 1LCK 15	
462	1lck	A	4	136	9e-31	0.68	1.00		P56==LCK== TYROSINE KINASE; 1LCK 7 CHAIN: A; 1LCK 8 TAIL PHOSPHOPEPTIDE TEGQ(PHOSPHO)YQPQPA; 1LCK 14 CHAIN: B; 1LCK 15	COMPLEX (KINASE/PEPTIDE)
462	1qcf	A	1	184	8e-44	0.29	1.00		HAEMATOPOETIC CELL KINASE (HCK); CHAIN: A;	TYROSINE KINASE TYROSINE KINASE-INHIBITOR COMPLEX, DOWN-REGULATED KINASE, 2 ORDERED ACTIVATION LOOP
462	1qgl	E	54	156	1.8e-24			97.37	GROWTH FACTOR RECEPTOR BINDING PROTEIN; CHAIN: E; SHC-DERIVED PEPTIDE; CHAIN: I;	HORMONE/GROWTH FACTOR GRB2-SH2; SIGNAL TRANSDUCTION, SH2 DOMAIN, PHOSPHOTYROSYL PEPTIDE, 2 COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE), HORMONE/GROWTH FACTOR
462	1qgl	E	58	148	1.8e-24	1.03	1.00		GROWTH FACTOR RECEPTOR BINDING PROTEIN; CHAIN: E; SHC-DERIVED PEPTIDE; CHAIN: I;	HORMONE/GROWTH FACTOR GRB2-SH2; SIGNAL TRANSDUCTION, SH2 DOMAIN, PHOSPHOTYROSYL PEPTIDE, 2 COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE), HORMONE/GROWTH FACTOR
462	1qpc	A	147	184	1.4e-08	0.06	-0.20		LCK KINASE; CHAIN: A;	TRANSFERASE ALPHA BETA FOLD
462	1sem	A	272	329	3.2e-20	-0.19	0.99		SEM-5; 1SEM 3 CHAIN: A, B; 1SEM 5 10-RESIDUE PROLINE-RICH PEPTIDE FROM MSOS 1SEM 8 CHAIN: C, D 1SEM 10	SIGNAL TRANSDUCTION PROTEIN SRC-HOMOLOGY 3 (SH3) DOMAIN, PEPTIDE-BINDING PROTEIN, 1SEM 18 2 GUANINE NUCLEOTIDE EXCHANGE FACTOR 1SEM 19
462	1sem	A	272	329	3.2e-20			58.54	SEM-5; 1SEM 3 CHAIN: A, B; 1SEM 5 10-RESIDUE PROLINE-RICH PEPTIDE FROM MSOS 1SEM 8 CHAIN: C, D 1SEM 10	SIGNAL TRANSDUCTION PROTEIN SRC-HOMOLOGY 3 (SH3) DOMAIN, PEPTIDE-BINDING PROTEIN, 1SEM 18 2 GUANINE NUCLEOTIDE EXCHANGE FACTOR 1SEM 19
462	1shf	A	273	328	1.4e-19	-0.31	1.00		PHOSPHOTRANSFERASE FYN PROTO-ONCOGENE TYROSINE KINASE (E.C.2.7.1.112) 1SHF 3 (SH3 DOMAIN) 1SHF 4	
462	1zfp	E	58	148	1.8e-27	0.92	1.00		GROWTH FACTOR RECEPTOR	COMPLEX (SIGNAL

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									BINDING PROTEIN; CHAIN: E; EPIDERMAL GROWTH FACTOR RECEPTOR-DERIVED PEPTIDE; CHAIN: I;	TRANSDUCTION/PEPTIDE) GRB2-SH2; 2-ABZ-GLU-TYR(PO3H2)-ILE-ASN-GLN-NH2, WITH 2-ABZ SIGNAL TRANSDUCTION, SH2 DOMAIN, PHOSPHOTYROSYL PEPTIDE, 2 COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE)
462	2shp	A	4	228	1.8e-10	0.16	0.49		SHP-2; CHAIN: A, B;	TYROSINE PHOSPHATASE SYP, SHPTP-2; TYROSINE PHOSPHATASE, INSULIN SIGNALING, SH2 PROTEIN
462	2shp	A	58	188	1.6e-14	0.14	1.00		SHP-2; CHAIN: A, B;	TYROSINE PHOSPHATASE SYP, SHPTP-2; TYROSINE PHOSPHATASE, INSULIN SIGNALING, SH2 PROTEIN
462	4hck		274	327	1.3e-17	0.18	1.00		HEMATOPOIETIC CELL KINASE; CHAIN: NULL;	TRANSFERASE HCK; SH3, PROTEIN TYROSINE KINASE, SIGNAL TRANSDUCTION, 2 TRANSFERASE
462	1ghu		56	149	5.4e-27			102.39	GRB2; CHAIN: NULL;	SRC HOMOLOGY 2 DOMAIN GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2; SRC HOMOLOGY 2 DOMAIN, GRB2, SH2
462	1ghu		58	148	5.4e-27	0.92	1.00		GRB2; CHAIN: NULL;	SRC HOMOLOGY 2 DOMAIN GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2; SRC HOMOLOGY 2 DOMAIN, GRB2, SH2
462	1gri	A	1	212	1.6e-40	0.85	1.00		GROWTH FACTOR BOUND PROTEIN 2; 1GRI 5 CHAIN: A, B; 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
462	1gri	A	1	212	1.6e-40			230.39	GROWTH FACTOR BOUND PROTEIN 2; 1GRI 5 CHAIN: A, B; 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
462	1gri	A	1	212	3.2e-37	0.76	1.00		GROWTH FACTOR BOUND PROTEIN 2; 1GRI 5 CHAIN: A, B; 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
462	1qgl	E	54	156	1.8e-24			101.17	GROWTH FACTOR RECEPTOR BINDING PROTEIN; CHAIN: E; SHC-DERIVED PEPTIDE; CHAIN: I;	HORMONE/GROWTH FACTOR GRB2-SH2; SIGNAL TRANSDUCTION, SH2 DOMAIN, PHOSPHOTYROSYL PEPTIDE, 2 COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE), HORMONE/GROWTH FACTOR
462	1qgl	E	58	148	1.8e-24	1.03	1.00		GROWTH FACTOR RECEPTOR BINDING PROTEIN; CHAIN: E; SHC-DERIVED PEPTIDE; CHAIN:	HORMONE/GROWTH FACTOR GRB2-SH2; SIGNAL TRANSDUCTION, SH2 DOMAIN, PHOSPHOTYROSYL PEPTIDE, 2 COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									I;	(SIGNAL TRANSDUCTION/PEPTIDE), HORMONE/GROWTH FACTOR
462	1zfp	E	58	148	1.8e-27	0.92	1.00		GROWTH FACTOR RECEPTOR BINDING PROTEIN; CHAIN: E; EPIDERMAL GROWTH FACTOR RECEPTOR-DERIVED PEPTIDE; CHAIN: I;	COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE) GRB2-SH2; 2-ABZ-GLU-TYR(PO3H2)-ILE-ASN-GLN-NH2, WITH 2-ABZ SIGNAL TRANSDUCTION, SH2 DOMAIN, PHOSPHOTYROSYL PEPTIDE, 2 COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE)
463	1efn	A	276	328	8e-19	-0.14	1.00		FYN TYROSINE KINASE; CHAIN: A, C; HIV-1 NEF PROTEIN; CHAIN: B, D;	COMPLEX (SH3 DOMAIN/VIRAL ENHANCER) SRC-HOMOLOGY 3 DOMAIN; COMPLEX (SH3 DOMAIN/VIRAL ENHANCER), PROTO-ONCOGENE, 2 TRANSFERASE, TYROSINE-PROTEIN KINASE, PHOSPHORYLATION, 3 AIDS, MYRISTYLATION, GTP-BINDING, ATP-BINDING, SH3 DOMAIN, 4 SH2 DOMAIN, PPII HELIX, PXXP MOTIF
463	1fmk		2	184	9.6e-45	0.22	1.00		TYROSINE-PROTEIN KINASE SRC; CHAIN: NULL;	PHOSPHOTRANSFERASE C-SRC, P60-SRC; SRC, TYROSINE KINASE, PHOSPHORYLATION, SH2, SH3, 2 PHOSPHOTYROSINE, PROTO-ONCOGENE, PHOSPHOTRANSFERASE
463	1fmk		273	327	4.8e-17	-0.02	1.00		TYROSINE-PROTEIN KINASE SRC; CHAIN: NULL;	PHOSPHOTRANSFERASE C-SRC, P60-SRC; SRC, TYROSINE KINASE, PHOSPHORYLATION, SH2, SH3, 2 PHOSPHOTYROSINE, PROTO-ONCOGENE, PHOSPHOTRANSFERASE
463	1fyn	A	273	328	1.4e-19	-0.08	1.00		PHOSPHOTRANSFERASE FYN; CHAIN: A; 3BP-2; CHAIN: B;	TRANSFERASE PROTO-ONCOGENE TYROSINE KINASE; PROTO-ONCOGENE, TRANSFERASE, TYROSINE-PROTEIN KINASE, 2 PHOSPHORYLATION, ATP-BINDING, MYRISTYLATION, SH3 DOMAIN, 3 COMPLEX (PHOSPHOTRANSFERASE/PEPTIDE)
463	1gbr	A	268	327	1.8e-17	-0.03	0.99		SIGNAL TRANSDUCTION PROTEIN GROWTH FACTOR	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									RECEPTOR-BOUND PROTEIN 2 (GRB2, N-TERMINAL 1GBR 3 SH3 DOMAIN) COMPLEXED WITH SOS-A PEPTIDE 1GBR 4 (NMR, 29 STRUCTURES) 1GBR 5	
463	1gfc		272	330	4.8e-21			63.10	ADAPTOR PROTEIN CONTAINING SH2 AND SH3 GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2 (GRB2) 1GFC 3 (C-TERMINAL SH3 DOMAIN) (NMR, MINIMIZED MEAN STRUCTURE) 1GFC 4	
463	1gfc		275	330	4.8e-21	-0.00	1.00		ADAPTOR PROTEIN CONTAINING SH2 AND SH3 GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2 (GRB2) 1GFC 3 (C-TERMINAL SH3 DOMAIN) (NMR, MINIMIZED MEAN STRUCTURE) 1GFC 4	
463	1ghu		56	149	5.4e-27			98.07	GRB2; CHAIN: NULL;	SRC HOMOLOGY 2 DOMAIN GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2; SRC HOMOLOGY 2 DOMAIN, GRB2, SH2
463	1ghu		58	148	5.4e-27	0.92	1.00		GRB2; CHAIN: NULL;	SRC HOMOLOGY 2 DOMAIN GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2; SRC HOMOLOGY 2 DOMAIN, GRB2, SH2
463	1ghu		58	148	9.6e-12	1.21	1.00		GRB2; CHAIN: NULL;	SRC HOMOLOGY 2 DOMAIN GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2; SRC HOMOLOGY 2 DOMAIN, GRB2, SH2
463	1gri	A	1	155	1.3e-28	0.81	1.00		GROWTH FACTOR BOUND PROTEIN 2; 1GRI 5 CHAIN: A, B; 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
463	1gri	A	1	216	1.3e-28			161.93	GROWTH FACTOR BOUND PROTEIN 2; 1GRI 5 CHAIN: A, B; 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
463	1gri	A	270	330	8e-22	0.33	1.00		GROWTH FACTOR BOUND PROTEIN 2; 1GRI 5 CHAIN: A, B; 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
463	1lck	A	2	149	9e-31			88.47	P56==LCK== TYROSINE KINASE;	COMPLEX (KINASE/PEPTIDE)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									1LCK 7 CHAIN: A; 1LCK 8 TAIL PHOSPHOPEPTIDE TEGQ(PHOSPHO)YQPQPA; 1LCK 14 CHAIN: B; 1LCK 15	
463	1lck	A	4	136	9e-31	0.68	1.00		P56=LCK=TYROSINE KINASE; 1LCK 7 CHAIN: A; 1LCK 8 TAIL PHOSPHOPEPTIDE TEGQ(PHOSPHO)YQPQPA; 1LCK 14 CHAIN: B; 1LCK 15	COMPLEX (KINASE/PEPTIDE)
463	1qcf	A	1	184	8e-44	0.29	1.00		HAEMATOPOETIC CELL KINASE (HCK); CHAIN: A;	TYROSINE KINASE TYROSINE KINASE-INHIBITOR COMPLEX, DOWN-REGULATED KINASE, 2 ORDERED ACTIVATION LOOP
463	1qg1	E	54	156	1.8e-24			97.37	GROWTH FACTOR RECEPTOR BINDING PROTEIN; CHAIN: E; SHC-DERIVED PEPTIDE; CHAIN: I;	HORMONE/GROWTH FACTOR GRB2-SH2; SIGNAL TRANSDUCTION, SH2 DOMAIN, PHOSPHOTYROSYL PEPTIDE, 2 COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE), HORMONE/GROWTH FACTOR
463	1qg1	E	58	148	1.8e-24	1.03	1.00		GROWTH FACTOR RECEPTOR BINDING PROTEIN; CHAIN: E; SHC-DERIVED PEPTIDE; CHAIN: I;	HORMONE/GROWTH FACTOR GRB2-SH2; SIGNAL TRANSDUCTION, SH2 DOMAIN, PHOSPHOTYROSYL PEPTIDE, 2 COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE), HORMONE/GROWTH FACTOR
463	1qpc	A	147	184	1.4e-08	0.06	-0.20		LCK KINASE; CHAIN: A;	TRANSFERASE ALPHA BETA FOLD
463	1sem	A	272	329	3.2e-20	-0.19	0.99		SEM-5; 1SEM 3 CHAIN: A, B; 1SEM 5 10-RESIDUE PROLINE-RICH PEPTIDE FROM MSOS 1SEM 8 CHAIN: C, D 1SEM 10	SIGNAL TRANSDUCTION PROTEIN SRC-HOMOLOGY 3 (SH3) DOMAIN, PEPTIDE-BINDING PROTEIN, 1SEM 18 2 GUANINE NUCLEOTIDE EXCHANGE FACTOR 1SEM 19
463	1sem	A	272	329	3.2e-20			58.54	SEM-5; 1SEM 3 CHAIN: A, B; 1SEM 5 10-RESIDUE PROLINE-RICH PEPTIDE FROM MSOS 1SEM 8 CHAIN: C, D 1SEM 10	SIGNAL TRANSDUCTION PROTEIN SRC-HOMOLOGY 3 (SH3) DOMAIN, PEPTIDE-BINDING PROTEIN, 1SEM 18 2 GUANINE NUCLEOTIDE EXCHANGE FACTOR 1SEM 19
463	1shf	A	273	328	1.4e-19	-0.31	1.00		PHOSPHOTRANSFERASE FYN PROTO-ONCOGENE TYROSINE KINASE (E.C.2.7.1.112) 1SHF 3 (SH3 DOMAIN) 1SHF 4	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
463	1zfp	E	58	148	1.8e-27	0.92	1.00		GROWTH FACTOR RECEPTOR BINDING PROTEIN; CHAIN: E; EPIDERMAL GROWTH FACTOR RECEPTOR-DERIVED PEPTIDE; CHAIN: I;	COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE) GRB2-SH2; 2-ABZ-GLU-TYR(PO3H2)-ILE-ASN-GLN-NH2, WITH 2-ABZ SIGNAL TRANSDUCTION, SH2 DOMAIN, PHOSPHOTYROSYL PEPTIDE, 2 COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE)
463	2shp	A	4	228	1.8e-10	0.16	0.49		SHP-2; CHAIN: A, B;	TYROSINE PHOSPHATASE SYP, SHPTP-2; TYROSINE PHOSPHATASE, INSULIN SIGNALING, SH2 PROTEIN
463	2shp	A	58	188	1.6e-14	0.14	1.00		SHP-2; CHAIN: A, B;	TYROSINE PHOSPHATASE SYP, SHPTP-2; TYROSINE PHOSPHATASE, INSULIN SIGNALING, SH2 PROTEIN
463	4hck		274	327	1.3e-17	0.18	1.00		HEMATOPOIETIC CELL KINASE; CHAIN: NULL;	TRANSFERASE HCK; SH3, PROTEIN TYROSINE KINASE, SIGNAL TRANSDUCTION, 2 TRANSFERASE
463	1ghu		56	149	5.4e-27			102.39	GRB2; CHAIN: NULL;	SRC HOMOLOGY 2 DOMAIN GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2; SRC HOMOLOGY 2 DOMAIN, GRB2, SH2
463	1ghu		58	148	5.4e-27	0.92	1.00		GRB2; CHAIN: NULL;	SRC HOMOLOGY 2 DOMAIN GROWTH FACTOR RECEPTOR-BOUND PROTEIN 2; SRC HOMOLOGY 2 DOMAIN, GRB2, SH2
463	1gri	A	1	212	1.6e-40	0.85	1.00		GROWTH FACTOR BOUND PROTEIN 2; 1GRI 5 CHAIN: A, B; 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
463	1gri	A	1	212	1.6e-40			230.39	GROWTH FACTOR BOUND PROTEIN 2; 1GRI 5 CHAIN: A, B; 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
463	1gri	A	1	212	3.2e-37	0.76	1.00		GROWTH FACTOR BOUND PROTEIN 2; 1GRI 5 CHAIN: A, B; 1GRI 6	SIGNAL TRANSDUCTION ADAPTOR SH2, SH3 1GRI 14
463	1qgl	E	54	156	1.8e-24			101.17	GROWTH FACTOR RECEPTOR BINDING PROTEIN; CHAIN: E; SHC-DERIVED PEPTIDE; CHAIN: I;	HORMONE/GROWTH FACTOR GRB2-SH2; SIGNAL TRANSDUCTION, SH2 DOMAIN, PHOSPHOTYROSYL PEPTIDE, 2 COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE), HORMONE/GROWTH FACTOR
463	1qgl	E	58	148	1.8e-24	1.03	1.00		GROWTH FACTOR RECEPTOR BINDING PROTEIN; CHAIN: E;	HORMONE/GROWTH FACTOR GRB2-SH2; SIGNAL TRANSDUCTION, SH2 DOMAIN,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									SHC-DERIVED PEPTIDE; CHAIN: I;	PHOSPHOTYROSYL PEPTIDE, 2 COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE), HORMONE/GROWTH FACTOR
463	1zfp	E	58	148	1.8e-27	0.92	1.00		GROWTH FACTOR RECEPTOR BINDING PROTEIN; CHAIN: E; EPIDERMAL GROWTH FACTOR RECEPTOR-DERIVED PEPTIDE; CHAIN: I;	COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE) GRB2-SH2; 2-ABZ-GLU-TYR(PO3H2)-ILE-ASN-GLN-NH2, WITH 2-ABZ SIGNAL TRANSDUCTION, SH2 DOMAIN, PHOSPHOTYROSYL PEPTIDE, 2 COMPLEX (SIGNAL TRANSDUCTION/PEPTIDE)
469	1pbw	A	329	504	3.6e-37	0.20	0.74		PHOSPHATIDYLINOSITOL 3-KINASE; CHAIN: A, B;	PHOSPHOTRANSFERASE RHOGAP DOMAIN; PHOSPHOTRANSFERASE, TPASE ACTIVATING PROTEIN, GAP, CDC42, 2 PHOSPHOINOSITIDE 3-KINASE, SH3 DOMAIN, SH2 DOMAIN, 3 SIGNAL TRANSDUCTION
469	1pbw	A	330	519	3.6e-37			79.76	PHOSPHATIDYLINOSITOL 3-KINASE; CHAIN: A, B;	PHOSPHOTRANSFERASE RHOGAP DOMAIN; PHOSPHOTRANSFERASE, TPASE ACTIVATING PROTEIN, GAP, CDC42, 2 PHOSPHOINOSITIDE 3-KINASE, SH3 DOMAIN, SH2 DOMAIN, 3 SIGNAL TRANSDUCTION
469	1pbw	A	368	516	6.4e-17	-0.01	0.34		PHOSPHATIDYLINOSITOL 3-KINASE; CHAIN: A, B;	PHOSPHOTRANSFERASE RHOGAP DOMAIN; PHOSPHOTRANSFERASE, TPASE ACTIVATING PROTEIN, GAP, CDC42, 2 PHOSPHOINOSITIDE 3-KINASE, SH3 DOMAIN, SH2 DOMAIN, 3 SIGNAL TRANSDUCTION
469	1pbw	B	329	504	1.8e-37	0.38	1.00		PHOSPHATIDYLINOSITOL 3-KINASE; CHAIN: A, B;	PHOSPHOTRANSFERASE RHOGAP DOMAIN; PHOSPHOTRANSFERASE; TPASE ACTIVATING PROTEIN, GAP, CDC42, 2 PHOSPHOINOSITIDE 3-KINASE, SH3 DOMAIN, SH2 DOMAIN, 3 SIGNAL TRANSDUCTION
469	1pbw	B	330	522	1.8e-37			77.88	PHOSPHATIDYLINOSITOL 3-KINASE; CHAIN: A, B;	PHOSPHOTRANSFERASE RHOGAP DOMAIN; PHOSPHOTRANSFERASE, TPASE ACTIVATING PROTEIN, GAP,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										CDC42, 2 PHOSPHOINOSITIDE 3-KINASE, SH3 DOMAIN, SH2 DOMAIN, 3 SIGNAL TRANSDUCTION
469	1pbw	B	368	516	6.4e-17	-0.04	0.39		PHOSPHATIDYLINOSITOL 3-KINASE; CHAIN: A, B;	PHOSPHOTRANSFERASE RHOGAP DOMAIN; PHOSPHOTRANSFERASE, TPASE ACTIVATING PROTEIN, GAP, CDC42, 2 PHOSPHOINOSITIDE 3-KINASE, SH3 DOMAIN, SH2 DOMAIN, 3 SIGNAL TRANSDUCTION
469	1rgp		313	502	3.6e-39			97.19	RHOGAP; CHAIN: NULL;	G-PROTEIN CDC42 GTPASE-ACTIVATING PROTEIN; G-PROTEIN, GAP, SIGNAL-TRANSDUCTION
469	1rgp		319	487	3.6e-39	0.13	1.00		RHOGAP; CHAIN: NULL;	G-PROTEIN CDC42 GTPASE-ACTIVATING PROTEIN; G-PROTEIN, GAP, SIGNAL-TRANSDUCTION
469	1rgp		367	523	1.6e-25	-0.11	0.82		RHOGAP; CHAIN: NULL;	G-PROTEIN CDC42 GTPASE-ACTIVATING PROTEIN; G-PROTEIN, GAP, SIGNAL-TRANSDUCTION
469	1tx4	A	316	523	1.8e-38			102.49	P50-RHOGAP; CHAIN: A; TRANSFORMING PROTEIN RHOA; CHAIN: B;	COMPLEX(GTPASE ACTIVATN/PROTO-ONCOGENE) GTPASE-ACTIVATING PROTEIN RHOGAP; COMPLEX (GTPASE ACTIVATION/PROTO-ONCOGENE), GTPASE, 2 TRANSITION STATE, GAP
469	1tx4	A	319	487	1.8e-38	0.24	1.00		P50-RHOGAP; CHAIN: A; TRANSFORMING PROTEIN RHOA; CHAIN: B;	COMPLEX(GTPASE ACTIVATN/PROTO-ONCOGENE) GTPASE-ACTIVATING PROTEIN RHOGAP; COMPLEX (GTPASE ACTIVATION/PROTO-ONCOGENE), GTPASE, 2 TRANSITION STATE, GAP
469	1tx4	A	367	523	1.6e-27	-0.09	0.83		P50-RHOGAP; CHAIN: A; TRANSFORMING PROTEIN RHOA; CHAIN: B;	COMPLEX(GTPASE ACTIVATN/PROTO-ONCOGENE) GTPASE-ACTIVATING PROTEIN RHOGAP; COMPLEX (GTPASE ACTIVATION/PROTO-ONCOGENE), GTPASE, 2 TRANSITION STATE, GAP
476	1fjg	I	270	396	3.2e-47	0.43	1.00		16S RIBOSOMAL RNA; CHAIN: A; FRAGMENT OF MESSENGER RNA; CHAIN: X; 30S RIBOSOMAL PROTEIN S2; CHAIN: B; 30S	RIBOSOME 30S RIBOSOMAL SUBUNIT, RIBOSOME, ANTIBIOTIC, STREPTOMYCIN, 2 SPECTINOMYCIN, PAROMOMYCIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									RIBOSOMAL PROTEIN S3; CHAIN: C; 30S RIBOSOMAL PROTEIN S4; CHAIN: D; 30S RIBOSOMAL PROTEIN S5; CHAIN: E; 30S RIBOSOMAL PROTEIN S6; CHAIN: F; 30S RIBOSOMAL PROTEIN S7; CHAIN: G; 30S RIBOSOMAL PROTEIN S8; CHAIN: H; 30S RIBOSOMAL PROTEIN S9; CHAIN: I; 30S RIBOSOMAL PROTEIN S10; CHAIN: J; 30S RIBOSOMAL PROTEIN S11; CHAIN: K; 30S RIBOSOMAL PROTEIN S12; CHAIN: L; 30S RIBOSOMAL PROTEIN S13; CHAIN: M; 30S RIBOSOMAL PROTEIN S14; CHAIN: N; 30S RIBOSOMAL PROTEIN S15; CHAIN: O; 30S RIBOSOMAL PROTEIN S16; CHAIN: P; 30S RIBOSOMAL PROTEIN S17; CHAIN: Q; 30S RIBOSOMAL PROTEIN S18; CHAIN: R; 30S RIBOSOMAL PROTEIN S19; CHAIN: S; 30S RIBOSOMAL PROTEIN S20; CHAIN: T; 30S RIBOSOMAL PROTEIN THX; CHAIN: V	
477	1a9n	A	28	107	1.8e-12	-0.28	0.33		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
477	1a9n	C	28	107	3.6e-12	-0.20	0.45		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
477	1d0b	A	20	147	6.4e-19	0.20	0.99		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
477	1d0b	A	5	125	1.6e-19	0.16	0.36		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
477	1dce	A	3	106	9.6e-15	0.28	0.37		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
477	1dce	A	50	148	3.2e-09	-0.15	0.71		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
477	1ds9	A	12	128	3.2e-14	-0.12	0.09		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
477	1yrg	A	28	119	7.2e-11	0.09	0.34		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI1, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE- 2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
477	2bnh		27	118	3.6e-11	-0.06	0.37		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
479	1fjg	R	65	130	1.6e-21	0.55	0.99		16S RIBOSOMAL RNA; CHAIN: A; FRAGMENT OF MESSENGER RNA; CHAIN: X; 30S RIBOSOMAL PROTEIN S2; CHAIN: B; 30S RIBOSOMAL PROTEIN S3; CHAIN: C; 30S RIBOSOMAL PROTEIN S4; CHAIN: D; 30S RIBOSOMAL PROTEIN S5; CHAIN: E; 30S RIBOSOMAL PROTEIN S6; CHAIN: F; 30S RIBOSOMAL PROTEIN S7; CHAIN: G; 30S RIBOSOMAL	RIBOSOME 30S RIBOSOMAL SUBUNIT, RIBOSOME, ANTIBIOTIC, STREPTOMYCIN, 2 SPECTINOMYCIN, PAROMOMYCIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									PROTEIN S8; CHAIN: H; 30S RIBOSOMAL PROTEIN S9; CHAIN: I; 30S RIBOSOMAL PROTEIN S10; CHAIN: J; 30S RIBOSOMAL PROTEIN S11; CHAIN: K; 30S RIBOSOMAL PROTEIN S12; CHAIN: L; 30S RIBOSOMAL PROTEIN S13; CHAIN: M; 30S RIBOSOMAL PROTEIN S14; CHAIN: N; 30S RIBOSOMAL PROTEIN S15; CHAIN: O; 30S RIBOSOMAL PROTEIN S16; CHAIN: P; 30S RIBOSOMAL PROTEIN S17; CHAIN: Q; 30S RIBOSOMAL PROTEIN S18; CHAIN: R; 30S RIBOSOMAL PROTEIN S19; CHAIN: S; 30S RIBOSOMAL PROTEIN S20; CHAIN: T; 30S RIBOSOMAL PROTEIN THX; CHAIN: V	
479	1fka	R	84	130	9.6e-18	-0.52	0.62		16S RIBOSOMAL RNA; CHAIN: A; 30S RIBOSOMAL PROTEIN S2; CHAIN: B; 30S RIBOSOMAL PROTEIN S3; CHAIN: C; 30S RIBOSOMAL PROTEIN S4; CHAIN: D; 30S RIBOSOMAL PROTEIN S5; CHAIN: E; 30S RIBOSOMAL PROTEIN S6; CHAIN: F; 30S RIBOSOMAL PROTEIN S7; CHAIN: G; 30S RIBOSOMAL PROTEIN S8; CHAIN: H; 30S RIBOSOMAL PROTEIN S9; CHAIN: I; 30S RIBOSOMAL PROTEIN S10; CHAIN: J; 30S RIBOSOMAL PROTEIN S11; CHAIN: K; 30S RIBOSOMAL PROTEIN S12; CHAIN: L; 30S RIBOSOMAL PROTEIN S13; CHAIN: M; 30S RIBOSOMAL PROTEIN S14;	RIBOSOME 30S RIBOSOMAL SUBUNIT, PROTEIN-RNA COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: N; 30S RIBOSOMAL PROTEIN S15; CHAIN: O; 30S RIBOSOMAL PROTEIN S16; CHAIN: P; 30S RIBOSOMAL PROTEIN S17; CHAIN: Q; 30S RIBOSOMAL PROTEIN S18; CHAIN: R; 30S RIBOSOMAL PROTEIN S19; CHAIN: S; 30S RIBOSOMAL PROTEIN S20; CHAIN: T	
479	1g1x	C	82	130	3.2e-13	0.38	1.00		30S RIBOSOMAL PROTEIN S6; CHAIN: A, F; 30S RIBOSOMAL PROTEIN S15; CHAIN: B, G; 30S RIBOSOMAL PROTEIN S18; CHAIN: C, H; 16S RIBOSOMAL RNA; CHAIN: D, I; 16S RIBOSOMAL RNA; CHAIN: E, J;	RIBOSOME RIBOSOMAL PROTEINS S15, S6, S18, S30 RIBOSOMAL SUBUNIT, RNA, 2 RIBOSOME
480	1fqv	B	20	114	1.1e-34	0.33	1.00		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
480	1fs1	B	20	114	4.8e-34	0.13	0.95		CYCLIN A/CDK2-ASSOCIATED P19; CHAIN: A, C; CYCLIN A/CDK2-ASSOCIATED P45; CHAIN: B, D;	LIGASE SKP2 F-BOX; SKP1; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
480	1fs2	B	20	114	6.4e-37	0.11	1.00		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
480	1vcb	B	20	115	7.2e-29	0.45	1.00		ELONGIN B; CHAIN: A, D, G, J; ELONGIN C; CHAIN: B, E, H, K; VHL; CHAIN: C, F, I, L;	TRANSCRIPTION TUMOR SUPPRESSOR, CANCER, UBIQUITIN, BETA SANDWICH, 2 TRANSCRIPTION, TRANSCRIPTIONAL ELONGATION
480	1vcb	B	20	115	7.2e-29			123.71	ELONGIN B; CHAIN: A, D, G, J;	TRANSCRIPTION TUMOR SUPPRESSOR,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									ELONGIN C; CHAIN: B, E, H, K; VHL; CHAIN: C, F, I, L;	CANCER, UBIQUITIN, BETA SANDWICH, 2 TRANSCRIPTION, TRANSCRIPTIONAL ELONGATION
480	1vcb	B	20	115	8e-28	0.45	1.00		ELONGIN B; CHAIN: A, D, G, J; ELONGIN C; CHAIN: B, E, H, K; VHL; CHAIN: C, F, I, L;	TRANSCRIPTION TUMOR SUPPRESSOR, CANCER, UBIQUITIN, BETA SANDWICH, 2 TRANSCRIPTION, TRANSCRIPTIONAL ELONGATION
482	1c3t	A	1	71	1.6e-26	0.76	1.00		1D8 UBIQUITIN; CHAIN: A;	DE NOVO PROTEIN PROTEIN DESIGN, HYDROPHOBIC CORE, PACKING, ROTAMERS, ROC, 2 UBIQUITIN, *DE NOVO* PROTEIN, UBIQUITIN
482	1c3t	A	1	76	9e-38	0.68	1.00		1D8 UBIQUITIN; CHAIN: A;	DE NOVO PROTEIN PROTEIN DESIGN, HYDROPHOBIC CORE, PACKING, ROTAMERS, ROC, 2 UBIQUITIN, DE NOVO PROTEIN, UBIQUITIN
482	1c3t	A	1	76	9e-38			106.98	1D8 UBIQUITIN; CHAIN: A;	DE NOVO PROTEIN PROTEIN DESIGN, HYDROPHOBIC CORE, PACKING, ROTAMERS, ROC, 2 UBIQUITIN, DE NOVO PROTEIN, UBIQUITIN
482	1c3t	A	77	152	8e-29	0.68	1.00		1D8 UBIQUITIN; CHAIN: A;	DE NOVO PROTEIN PROTEIN DESIGN, HYDROPHOBIC CORE, PACKING, ROTAMERS, ROC, 2 UBIQUITIN, DE NOVO PROTEIN, UBIQUITIN
482	1c3t	A	77	152	9e-38	0.68	1.00		1D8 UBIQUITIN; CHAIN: A;	DE NOVO PROTEIN PROTEIN DESIGN, HYDROPHOBIC CORE, PACKING, ROTAMERS, ROC, 2 UBIQUITIN, DE NOVO PROTEIN, UBIQUITIN
482	1tbe	B	1	71	3.2e-28	1.19	1.00		UBIQUITIN TETRAUBIQUITIN 1TBE 3	
482	1tbe	B	1	72	5.4e-35	0.97	1.00		UBIQUITIN TETRAUBIQUITIN 1TBE 3	
482	1tbe	B	1	72	5.4e-35			102.16	UBIQUITIN TETRAUBIQUITIN 1TBE 3	
482	1tbe	B	77	148	1.4e-27	0.97	1.00		UBIQUITIN TETRAUBIQUITIN 1TBE 3	
482	1tbe	B	77	148	5.4e-35	0.97	1.00		UBIQUITIN TETRAUBIQUITIN 1TBE 3	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
482	1ubi		1	71	3.2e-28	1.28	1.00		CHROMOSOMAL PROTEIN UBIQUITIN 1UBI 3	
482	1ubi		1	76	7.2e-36	1.07	1.00		CHROMOSOMAL PROTEIN UBIQUITIN 1UBI 3	
482	1ubi		1	76	7.2e-36			110.47	CHROMOSOMAL PROTEIN UBIQUITIN 1UBI 3	
482	1ubi		77	152	1.6e-30	1.07	1.00		CHROMOSOMAL PROTEIN UBIQUITIN 1UBI 3	
482	1ubi		77	152	7.2e-36	1.07	1.00		CHROMOSOMAL PROTEIN UBIQUITIN 1UBI 3	
482	1ud7	A	1	71	8e-27	0.76	1.00		UBIQUITIN CORE MUTANT 1D7; CHAIN: A;	UBIQUITIN UBIQUITIN, DESIGNED CORE MUTANT
482	1ud7	A	1	76	1.3e-35	0.96	1.00		UBIQUITIN CORE MUTANT 1D7; CHAIN: A;	UBIQUITIN UBIQUITIN, DESIGNED CORE MUTANT
482	1ud7	A	1	76	1.3e-35			106.97	UBIQUITIN CORE MUTANT 1D7; CHAIN: A;	UBIQUITIN UBIQUITIN, DESIGNED CORE MUTANT
482	1ud7	A	77	152	1.3e-35	0.96	1.00		UBIQUITIN CORE MUTANT 1D7; CHAIN: A;	UBIQUITIN UBIQUITIN, DESIGNED CORE MUTANT
482	1ud7	A	77	152	3.2e-29	0.96	1.00		UBIQUITIN CORE MUTANT 1D7; CHAIN: A;	UBIQUITIN UBIQUITIN, DESIGNED CORE MUTANT
483	1c4o	A	197	385	1.1e-16	0.09	0.89		DNA NUCLEOTIDE EXCISION REPAIR ENZYME UVRB; CHAIN: A;	REPLICATION DNA NUCLEOTIDE EXCISION REPAIR, UVRABC, HELICASE, 2 HYPERTHERMOSTABLE PROTEIN
483	1d2m	A	197	385	1.4e-16	-0.10	0.77		EXCINUCLEASE ABC SUBUNIT B; CHAIN: A;	HYDROLASE UVRB; MULTIDOMAIN PROTEIN
483	1d9x	A	134	378	3.6e-40	-0.16	0.78		EXCINUCLEASE UVRABC COMPONENT UVRB; CHAIN: A;	GENE REGULATION APO PROTEIN
483	1d9x	A	167	395	1.4e-18	-0.19	0.98		EXCINUCLEASE UVRABC COMPONENT UVRB; CHAIN: A;	GENE REGULATION APO PROTEIN
483	1fuk	A	242	401	3.2e-45	0.39	1.00		EUKARYOTIC INITIATION FACTOR 4A; CHAIN: A;	TRANSLATION YEAST INITIATION FACTOR 4A, EIF4A; HELICASE, INITIATION FACTOR 4A, DEAD-BOX PROTEIN
483	1fim	A	8	226	1.3e-57	0.93	1.00	-	YEAST INITIATION FACTOR 4A; CHAIN: A, B;	TRANSLATION EUKARYOTIC INITIATION FACTOR 4A; IF4A, HELICASE, DEAD-BOX PROTEIN
483	1fim	B	8	401	0	0.62	1.00		YEAST INITIATION FACTOR 4A;	TRANSLATION EUKARYOTIC INITIATION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: A, B;	FACTOR 4A; IF4A, HELICASE, DEAD-BOX PROTEIN
483	1hei	A	262	377	1.6e-09	0.26	0.01		HCV HELICASE; CHAIN: A, B;	HELICASE HELICASE, RNA, HEPATITIS, HCV, ATPASE, NTPASE
483	1hei	B	262	377	1.6e-09	-0.24	0.04		HCV HELICASE; CHAIN: A, B;	HELICASE HELICASE, RNA, HEPATITIS, HCV, ATPASE, NTPASE
483	1qde	A	8	225	3.2e-54	0.54	1.00		TRANSLATION INITIATION FACTOR 4A; CHAIN: A;	GENE REGULATION EIF4A; TRANSLATION INITIATION, SACCHAROMYCES CEREVISIAE, DEAD BOX 2 PROTEIN FAMILY
483	8ohm		41	368	9e-58	-0.22	0.04		RNA HELICASE; CHAIN: NULL	HELICASE RNA HELICASE, HEPATITIS C VIRUS, HCV, UNWINDING MECHANISM
490	1b6e		99	220	5.4e-27			84.94	CD94; CHAIN: NULL;	NK CELL NK CELL, RECEPTOR, C-TYPE LECTIN, C-TYPE LECTIN-LIKE, NKD
490	1bj3	A	101	210	1.6e-34			54.46	COAGULATION FACTOR IX-BINDING PROTEIN A; CHAIN: A; COAGULATION FACTOR IX-BINDING PROTEIN B; CHAIN: B;	COLLAGEN BINDING PROTEIN IX-BP; IX-BP; COAGULATION FACTOR IX-BINDING, HETERODIMER, VENOM, HABU 2 SNAKE, C-TYPE LECTIN SUPERFAMILY, COLLAGEN BINDING PROTEIN
490	1bj3	A	102	217	1.6e-34	0.23	0.29		COAGULATION FACTOR IX-BINDING PROTEIN A; CHAIN: A; COAGULATION FACTOR IX-BINDING PROTEIN B; CHAIN: B;	COLLAGEN BINDING PROTEIN IX-BP; IX-BP; COAGULATION FACTOR IX-BINDING, HETERODIMER, VENOM, HABU 2 SNAKE, C-TYPE LECTIN SUPERFAMILY, COLLAGEN BINDING PROTEIN
490	1c3a	B	101	220	6.4e-33	0.23	0.95		FLAVOCETIN-A: ALPHA SUBUNIT; CHAIN: A; FLAVOCETIN-A: BETA SUBUNIT; CHAIN: B	MEMBRANE PROTEIN C-TYPE LECTIN-LIKE DOMAINS
490	1dv8	A	102	217	6.4e-33	0.69	1.00		ASIALOGLYCOPROTEIN RECEPTOR 1; CHAIN: A;	SIGNALING PROTEIN HEPATIC LECTIN H1; C-TYPE LECTIN CRD
490	1e87	A	100	219	1.1e-27	0.52	0.86		EARLY ACTIVATION ANTIGEN CD69; CHAIN: A;	HEMATOPOIETIC CELL RECEPTOR ACTIVATION INDUCER MOLECULE (AIM), EA 1, HEMATOPOIETIC CELL RECEPTOR, LEUCOCYTE, C-TYPE LECTIN-LIKE, 2 NKD, KLR
490	1ixx	A	101	218	6.4e-32			52.61	COAGULATION FACTORS IX/X-BINDING PROTEIN; CHAIN: A, B,	COAGULATION FACTOR BINDING IX/X-BP COAGULATION FACTOR BINDING, C-

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									C, D, E, F;	TYPE LECTIN, GLA-DOMAIN 2 BINDING, C-TYPE CRD MOTIF, LOOP EXCHANGED DIMER
490	1ixx	A	102	217	6.4e-32	0.00	0.22		COAGULATION FACTORS IX/X-BINDING PROTEIN; CHAIN: A, B, C, D, E, F;	COAGULATION FACTOR BINDING IX/X-BP COAGULATION FACTOR BINDING, C-TYPE LECTIN, GLA-DOMAIN 2 BINDING, C-TYPE CRD MOTIF, LOOP EXCHANGED DIMER
490	1ixx	B	101	220	1.6e-33			59.42	COAGULATION FACTORS IX/X-BINDING PROTEIN; CHAIN: A, B, C, D, E, F;	COAGULATION FACTOR BINDING IX/X-BP COAGULATION FACTOR BINDING, C-TYPE LECTIN, GLA-DOMAIN 2 BINDING, C-TYPE CRD MOTIF, LOOP EXCHANGED DIMER
490	1ixx	B	102	220	1.6e-33	0.33	0.76		COAGULATION FACTORS IX/X-BINDING PROTEIN; CHAIN: A, B, C, D, E, F;	COAGULATION FACTOR BINDING IX/X-BP COAGULATION FACTOR BINDING, C-TYPE LECTIN, GLA-DOMAIN 2 BINDING, C-TYPE CRD MOTIF, LOOP EXCHANGED DIMER
490	1lit		102	219	4.8e-34	0.61	1.00		LITHOSTATHINE; CHAIN: NULL	PANCREATIC STONE INHIBITOR PANCREATIC STONE INHIBITOR, LECTIN
490	1lit		102	220	4.8e-34			54.29	LITHOSTATHINE; CHAIN: NULL	PANCREATIC STONE INHIBITOR PANCREATIC STONE INHIBITOR, LECTIN
490	1qdd	A	89	220	8e-35			59.45	LITHOSTATHINE; CHAIN: A;	METAL BINDING PROTEIN PANCREATIC STONE PROTEIN, PSP; PANCREATIC STONE INHIBITOR, LITHOSTATHINE
490	1qdd	A	91	219	8e-35	0.44	0.72		LITHOSTATHINE; CHAIN: A;	METAL BINDING PROTEIN PANCREATIC STONE PROTEIN, PSP; PANCREATIC STONE INHIBITOR, LITHOSTATHINE
490	1qo3	C	97	219	7.2e-29	0.35	0.74		MHC CLASS I H-2DD HEAVY CHAIN; CHAIN: A; BETA-2-MICROGLOBULIN; CHAIN: B; HIV ENVELOPE GLYCOPROTEIN 120 PEPTIDE; CHAIN: P; LY49A; CHAIN: C, D;	COMPLEX (NK RECEPTOR/MHC CLASS I) H-2 CLASS I HISTOCOMPATIBILITY ANTIGEN, B2M; NK-CELL SURFACE GLYCOPROTEIN YE1/48, NK CELL, INHIBITORY RECEPTOR, MHC-I, C-TYPE LECTIN-LIKE, 2 HISTOCOMPATIBILITY, B2M, LY49, LY-49
490	2afp	A	96	216	4.8e-30	0.32	0.65		SEA RAVEN TYPE II ANTIFREEZE PROTEIN; CHAIN: A;	ANTIFREEZE PROTEIN RECOMBINANT SEA RAVEN PROTEIN, SOLUTION BACKBONE FOLD, C-2 TYPE LECTIN,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										ANTIFREEZE PROTEIN
492	1c4o	A	274	461	1.4e-18	-0.19	0.24		DNA NUCLEOTIDE EXCISION REPAIR ENZYME UVRB; CHAIN: A;	REPLICATION DNA NUCLEOTIDE EXCISION REPAIR, UVRABC, HELICASE, 2 HYPERTHERMOSTABLE PROTEIN
492	1d2m	A	274	461	1.1e-18	0.21	0.54		EXCINUCLEASE ABC SUBUNIT B; CHAIN: A;	HYDROLASE UVRB; MULTIDOMAIN PROTEIN
492	1d9x	A	222	481	3.6e-43	-0.39	0.16		EXCINUCLEASE UVRABC COMPONENT UVRB; CHAIN: A;	GENE REGULATION APO PROTEIN
492	1d9x	A	274	469	6.4e-22	-0.03	0.59		EXCINUCLEASE UVRABC COMPONENT UVRB; CHAIN: A;	GENE REGULATION APO PROTEIN
492	1fuk	A	312	478	3.2e-50	0.64	1.00		EUKARYOTIC INITIATION FACTOR 4A; CHAIN: A;	TRANSLATION YEAST INITIATION FACTOR 4A, EIF4A; HELICASE, INITIATION FACTOR 4A, DEAD-BOX PROTEIN
492	1fuu	A	111	304	1.4e-54	0.53	1.00		YEAST INITIATION FACTOR 4A; CHAIN: A, B;	TRANSLATION EUKARYOTIC INITIATION FACTOR 4A; IF4A, HELICASE, DEAD-BOX PROTEIN
492	1fuu	B	111	478	0	0.54	1.00		YEAST INITIATION FACTOR 4A; CHAIN: A, B;	TRANSLATION EUKARYOTIC INITIATION FACTOR 4A; IF4A, HELICASE, DEAD-BOX PROTEIN
492	1qde	A	111	302	3.2e-52	0.58	1.00		TRANSLATION INITIATION FACTOR 4A; CHAIN: A;	GENE REGULATION EIF4A; TRANSLATION INITIATION, SACCHAROMYCES CEREVISIAE, DEAD BOX 2 PROTEIN FAMILY
493	1faq		462	487	0.0036	-0.71	0.06		RAF-1; CHAIN: NULL;	SERINE/THREONINE PROTEIN KINASE TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, 2 PROTO-ONCOGENE, ZINC, ATP-BINDING, PHORBOL-ESTER BINDING
493	1ptq		462	497	0.0072	-0.06	0.03		PROTEIN KINASE C DELTA TYPE; 1PTQ 4	PHOSPHOTRANSFERASE
494	1a1h	A	112	220	3.2e-27	0.01	0.94		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
494	1alh	A	140	222	3.2e-27			77.15	QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
494	1alh	A	140	238	1.4e-20	0.05	0.30		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
494	1mey	C	111	192	4.8e-39	-0.40	0.68		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
494	1mey	C	139	220	4.8e-39	0.33	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
494	1mey	C	139	221	4.8e-39			88.65	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
494	1mey	C	195	239	9.6e-22	-0.11	0.42		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
494	1mey	C	82	164	6.4e-42	-0.21	0.81		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
494	1mey	G	110	136	1.6e-12	-0.38	0.04		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
494	1mey	G	193	220	1.6e-13	0.11	0.98		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
494	1tf3	A	112	224	3.2e-16	-0.16	0.39		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
494	1tf3	A	139	224	3.2e-16			55.53	TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
494	1tf3	A	185	228	6.4e-12	0.23	0.45		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
494	1tf6	A	113	221	3.6e-25	-0.26	0.03		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
494	1tf6	A	81	254	4.8e-28			68.71	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
494	1ubd	C	113	221	1.8e-31			79.21	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1;

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									INITIATOR ELEMENT DNA; CHAIN: A, B;	TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
494	1ubd	C	116	221	1.8e-31	-0.02	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
494	1ubd	C	119	238	3.2e-23	-0.40	0.49		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
494	1ubd	C	58	220	8e-26	-0.35	0.31		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
494	2adr		140	222	1.6e-12	-0.02	0.69		ADR1; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADR1, ZINC FINGER, NMR
494	2adr		168	226	1.6e-12			55.92	ADR1; CHAIN: NULL;	TRANSCRIPTION REGULATION TRANSCRIPTION REGULATION, ADR1, ZINC FINGER, NMR
494	2gli	A	113	221	9e-29	-0.08	0.87		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
494	2gli	A	123	235	4.8e-20	-0.25	0.16		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
494	2gli	A	185	239	8e-15	-0.22	0.24		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
494	2gli	A	40	222	3.2e-26	-0.24	0.06		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
494	2gli	A	74	222	9e-29			76.87	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
499	1erj	A	132	434	9.6e-61	-0.02	0.96		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
499	1erj	A	214	459	4.8e-44	-0.15	0.04		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
499	1erj	A	35	432	3.6e-20	0.15	0.96		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
499	1erj	A	42	353	1.4e-64	0.27	1.00		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
499	1got	B	127	432	1.6e-62	0.12	0.88		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
499	1got	B	34	350	9.6e-74	0.33	1.00		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
499	1got	B	34	371	9.6e-74			69.42	GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										TRANSDUCTION
500	1bc9		130	319	1.1e-65			98.84	CYTOHESIN-1; CHAIN: NULL;	EXCHANGE FACTOR B2-1, SEC7 HOMOLOG B2-1; EXCHANGE FACTOR, INTEGRIN BINDING PROTEIN
500	1bc9		134	311	1.1e-65	0.58	1.00		CYTOHESIN-1; CHAIN: NULL;	EXCHANGE FACTOR B2-1, SEC7 HOMOLOG B2-1; EXCHANGE FACTOR, INTEGRIN BINDING PROTEIN
500	1fqv	A	69	113	1.6e-10	-0.30	0.03		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
500	1fs1	A	71	108	1.1e-08	-0.43	0.29		CYCLIN A/CDK2-ASSOCIATED P19; CHAIN: A, C; CYCLIN A/CDK2-ASSOCIATED P45; CHAIN: B, D;	LIGASE SKP2 F-BOX; SKP1; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
500	1fs2	A	69	147	1.4e-12	-0.20	0.15		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
500	1pbv		126	319	3.2e-65			107.50	ARNO; CHAIN: NULL;	EXCHANGE FACTOR ARF NUCLEOTIDE-BINDING SITE OPENER; EXCHANGE FACTOR, SEC7, ARNO, ARF FUNCTIONAL CLASS: GUANINE 2 NUCLEOTIDE EXCHANGE FACTOR
500	1pbv		127	311	3.2e-65	0.16	1.00		ARNO; CHAIN: NULL;	EXCHANGE FACTOR ARF NUCLEOTIDE-BINDING SITE OPENER; EXCHANGE FACTOR, SEC7, ARNO, ARF FUNCTIONAL CLASS: GUANINE 2 NUCLEOTIDE EXCHANGE FACTOR
503	1a1h	A	145	223	4.8e-25	-0.28	0.09		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
503	1alh	A	339	397	4.8e-23	-0.19	0.80		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
503	1alh	A	339	398	1.1e-25	-0.20	0.53		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
503	1mey	C	144	223	1.6e-43	0.03	0.43		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
503	1mey	C	171	251	1.6e-46	-0.04	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
503	1mey	C	198	279	3.2e-50	0.39	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
503	1mey	C	226	307	8e-51	0.27	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
503	1mey	C	226	308	4.8e-51			93.44	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
503	1mey	C	254	335	4.8e-51	0.05	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
503	1mey	C	282	363	4.8e-51	-0.08	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
503	1mey	C	310	391	1.6e-51	-0.36	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
503	1mey	C	338	398	1.1e-37	-0.47	0.96		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
503	1tf6	A	172	316	4.8e-37	-0.06	0.75		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
503	1tf6	A	196	372	3.6e-71			95.64	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
503	1tf6	A	255	393	1.6e-36	-0.22	0.94		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
503	1ubd	C	152	251	1.6e-30	0.20	0.52		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
503	1ubd	C	179	279	1.6e-35	0.18	0.99		YY1; CHAIN: C; ADENO-	COMPLEX (TRANSCRIPTION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
503	1ubd	C	180	279	9e-41	0.05	0.69		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
503	1ubd	C	200	308	3.6e-53			77.59	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
503	1ubd	C	203	307	5.4e-50	0.14	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
503	1ubd	C	224	364	3.6e-53	-0.36	0.33		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
503	1ubd	C	280	391	1.8e-47	-0.35	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
503	1ubd	C	290	391	1.3e-34	-0.09	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
503	1ubd	C	318	397	4.8e-28	-0.48	0.78		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
503	2gli	A	162	281	1.1e-39	0.01	0.93		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
503	2gli	A	179	306	3.2e-34	0.35	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
503	2gli	A	198	337	1.4e-66			78.17	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
503	2gli	A	198	337	3.6e-64	0.12	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
503	2gli	A	226	364	1.4e-66	0.04	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
503	2gli	A	254	392	3.6e-62	0.05	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
503	2gli	A	262	390	4.8e-33	-0.04	0.95		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
503	2gli	A	290	397	1.1e-27	-0.28	0.43		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
506	1a7t	A	31	197	4.8e-18	0.04	-0.15		METALLO-BETA-LACTAMASE; CHAIN: A, B;	HYDROLASE CLASS B BETA-LACTAMASE; HYDROLASE (BETA-LACTAMASE), METALLO BETA-LACTAMASE, ZINC
506	1a7t	A	37	196	3.6e-19	0.30	0.78		METALLO-BETA-LACTAMASE; CHAIN: A, B;	HYDROLASE CLASS B BETA-LACTAMASE; HYDROLASE (BETA-LACTAMASE), METALLO BETA-LACTAMASE, ZINC
506	1dd6	A	47	197	7.2e-25	-0.03	0.42		IMP-1 METALLO BETA-LACTAMASE; CHAIN: A, B;	HYDROLASE METALLO BETA-LACTAMASE INHIBITOR, MERCAPTOCARBOXYLATE 2 INHIBITOR, IMP-1 METALLO BETA-LACTAMASE
506	1e5d	A	2	243	3.2e-28	0.15	-0.18		RUBREDOXIN:OXYGEN OXIDOREDUCTASE; CHAIN: A, B	OXIDOREDUCTASE OXIDOREDUCTASE, OXYGENREDUCTASE, DIIRON-CENTRE, 2 FLAVOPROTEINS, LACTAMASE-FOLD
506	1qh5	A	24	254	4.8e-49			81.91	HYDROXYACYLGLUTATHIONE HYDROLASE; CHAIN: A, B;	HYDROLASE GLYOXALASE II; METALLO-HYDROLASE
506	1qh5	A	33	253	4.8e-49	0.37	1.00		HYDROXYACYLGLUTATHIONE HYDROLASE; CHAIN: A, B;	HYDROLASE GLYOXALASE II; METALLO-HYDROLASE
506	1sml	A	37	201	1.1e-24	0.44	0.88		PENICILLINASE; CHAIN: A;	HYDROLASE METALLO-BETA-LACTAMASE, ANTIBIOTIC RESISTANCE, BINUCLEAR 2 ZINC, HYDROLASE
506	2bc2	A	24	198	6.4e-20	0.07	-0.12		METALLO BETA-LACTAMASE II; CHAIN: A, B;	HYDROLASE HYDROLASE, BETA-LACTAMASE, ANTIBIOTIC, METALLOENZYME
506	2bc2	A	37	197	3.6e-27	0.20	0.18		METALLO BETA-LACTAMASE II; CHAIN: A, B;	HYDROLASE HYDROLASE, BETA-LACTAMASE, ANTIBIOTIC, METALLOENZYME
508	1dm9	A	107	170	3.2e-05	-0.60	0.05		HYPOTHETICAL 15.5 KD PROTEIN IN MRCA-PCKA CHAIN: A, B	STRUCTURAL GENOMICS HEAT SHOCK PROTEINS, PROTEIN-RNA INTERACTIONS, RIBOSOME, 2 STRUCTURAL GENOMICS
508	1fjg	D	20	171	1.1e-43	-0.31	0.18		16S RIBOSOMAL RNA; CHAIN: A; FRAGMENT OF MESSENGER	RIBOSOME 30S RIBOSOMAL SUBUNIT, RIBOSOME, ANTIBIOTIC,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									RNA; CHAIN: X; 30S RIBOSOMAL PROTEIN S2; CHAIN: B; 30S RIBOSOMAL PROTEIN S3; CHAIN: C; 30S RIBOSOMAL PROTEIN S4; CHAIN: D; 30S RIBOSOMAL PROTEIN S5; CHAIN: E; 30S RIBOSOMAL PROTEIN S6; CHAIN: F; 30S RIBOSOMAL PROTEIN S7; CHAIN: G; 30S RIBOSOMAL PROTEIN S8; CHAIN: H; 30S RIBOSOMAL PROTEIN S9; CHAIN: I; 30S RIBOSOMAL PROTEIN S10; CHAIN: J; 30S RIBOSOMAL PROTEIN S11; CHAIN: K; 30S RIBOSOMAL PROTEIN S12; CHAIN: L; 30S RIBOSOMAL PROTEIN S13; CHAIN: M; 30S RIBOSOMAL PROTEIN S14; CHAIN: N; 30S RIBOSOMAL PROTEIN S15; CHAIN: O; 30S RIBOSOMAL PROTEIN S16; CHAIN: P; 30S RIBOSOMAL PROTEIN S17; CHAIN: Q; 30S RIBOSOMAL PROTEIN S18; CHAIN: R; 30S RIBOSOMAL PROTEIN S19; CHAIN: S; 30S RIBOSOMAL PROTEIN S20; CHAIN: T; 30S RIBOSOMAL PROTEIN THX; CHAIN: V	STREPTOMYCIN, 2 SPECTINOMYCIN, PAROMOMYCIN
508	1fka	D	54	171	8e-36	-0.47	0.15		16S RIBOSOMAL RNA; CHAIN: A; 30S RIBOSOMAL PROTEIN S2; CHAIN: B; 30S RIBOSOMAL PROTEIN S3; CHAIN: C; 30S RIBOSOMAL PROTEIN S4; CHAIN: D; 30S RIBOSOMAL PROTEIN S5; CHAIN: E; 30S RIBOSOMAL PROTEIN S6; CHAIN: F; 30S RIBOSOMAL PROTEIN S7; CHAIN: G; 30S RIBOSOMAL PROTEIN S8;	RIBOSOME 30S RIBOSOMAL SUBUNIT, PROTEIN-RNA COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: H; 30S RIBOSOMAL PROTEIN S9; CHAIN: I; 30S RIBOSOMAL PROTEIN S10; CHAIN: J; 30S RIBOSOMAL PROTEIN S11; CHAIN: K; 30S RIBOSOMAL PROTEIN S12; CHAIN: L; 30S RIBOSOMAL PROTEIN S13; CHAIN: M; 30S RIBOSOMAL PROTEIN S14; CHAIN: N; 30S RIBOSOMAL PROTEIN S15; CHAIN: O; 30S RIBOSOMAL PROTEIN S16; CHAIN: P; 30S RIBOSOMAL PROTEIN S17; CHAIN: Q; 30S RIBOSOMAL PROTEIN S18; CHAIN: R; 30S RIBOSOMAL PROTEIN S19; CHAIN: S; 30S RIBOSOMAL PROTEIN S20; CHAIN: T	
508	1qd7	C	55	171	3.2e-33	-0.43	0.21		CENTRAL FRAGMENT OF 16 S RNA; CHAIN: A; END FRAGMENT OF 16 S RNA; CHAIN: B; S4 RIBOSOMAL PROTEIN; CHAIN: C; S5 RIBOSOMAL PROTEIN; CHAIN: D; S6 RIBOSOMAL PROTEIN; CHAIN: E; S7 RIBOSOMAL PROTEIN; CHAIN: F; S8 RIBOSOMAL PROTEIN; CHAIN: G; S15 RIBOSOMAL PROTEIN; CHAIN: H; S17 RIBOSOMAL PROTEIN; CHAIN: I; S20 RIBOSOMAL PROTEIN; CHAIN: J	RIBOSOME 30S RIBOSOMAL SUBUNIT, LOW RESOLUTION MODEL
510	1hlg	A	113	158	0.009	-0.74	0.04		LIPASE, GASTRIC; CHAIN: A, B;	HYDROLASE LIPASE
512	1pbw	A	29	227	1.4e-24	0.31	1.00		PHOSPHATIDYLINOSITOL 3-KINASE; CHAIN: A, B;	PHOSPHOTRANSFERASE RHOGAP DOMAIN; PHOSPHOTRANSFERASE, TPASE ACTIVATING PROTEIN, GAP,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										CDC42, 2 PHOSPHOINOSITIDE 3-KINASE, SH3 DOMAIN, SH2 DOMAIN, 3 SIGNAL TRANSDUCTION
512	1pbw	A	29	229	5.4e-43			83.39	PHOSPHATIDYLINOSITOL 3-KINASE; CHAIN: A, B;	PHOSPHOTRANSFERASE RHOGAP DOMAIN; PHOSPHOTRANSFERASE, TPASE ACTIVATING PROTEIN, GAP, CDC42, 2 PHOSPHOINOSITIDE 3-KINASE, SH3 DOMAIN, SH2 DOMAIN, 3 SIGNAL TRANSDUCTION
512	1pbw	A	34	229	5.4e-43	0.47	1.00		PHOSPHATIDYLINOSITOL 3-KINASE; CHAIN: A, B;	PHOSPHOTRANSFERASE RHOGAP DOMAIN; PHOSPHOTRANSFERASE, TPASE ACTIVATING PROTEIN, GAP, CDC42, 2 PHOSPHOINOSITIDE 3-KINASE, SH3 DOMAIN, SH2 DOMAIN, 3 SIGNAL TRANSDUCTION
512	1pbw	B	29	227	1.4e-24	0.36	1.00		PHOSPHATIDYLINOSITOL 3-KINASE; CHAIN: A, B;	PHOSPHOTRANSFERASE RHOGAP DOMAIN; PHOSPHOTRANSFERASE, TPASE ACTIVATING PROTEIN, GAP, CDC42, 2 PHOSPHOINOSITIDE 3-KINASE, SH3 DOMAIN, SH2 DOMAIN, 3 SIGNAL TRANSDUCTION
512	1pbw	B	29	235	1.8e-44	0.50	1.00		PHOSPHATIDYLINOSITOL 3-KINASE; CHAIN: A, B;	PHOSPHOTRANSFERASE RHOGAP DOMAIN; PHOSPHOTRANSFERASE, TPASE ACTIVATING PROTEIN, GAP, CDC42, 2 PHOSPHOINOSITIDE 3-KINASE, SH3 DOMAIN, SH2 DOMAIN, 3 SIGNAL TRANSDUCTION
512	1pbw	B	29	235	1.8e-44			84.44	PHOSPHATIDYLINOSITOL 3-KINASE; CHAIN: A, B;	PHOSPHOTRANSFERASE RHOGAP DOMAIN; PHOSPHOTRANSFERASE, TPASE ACTIVATING PROTEIN, GAP, CDC42, 2 PHOSPHOINOSITIDE 3-KINASE, SH3 DOMAIN, SH2 DOMAIN, 3 SIGNAL TRANSDUCTION
512	1rgp		16	223	3.6e-51			112.26	RHOGAP; CHAIN: NULL;	G-PROTEIN CDC42 GTPASE-ACTIVATING PROTEIN; G-PROTEIN, GAP, SIGNAL-TRANSDUCTION
512	1rgp		16	234	1.6e-39	0.45	1.00		RHOGAP; CHAIN: NULL;	G-PROTEIN CDC42 GTPASE-ACTIVATING PROTEIN; G-PROTEIN, GAP, SIGNAL-

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										TRANSDUCTION
512	lrgp		16	234	3.6e-51	0.65	1.00		RHOGAP; CHAIN: NULL;	G-PROTEIN CDC42 GTPASE-ACTIVATING PROTEIN; G-PROTEIN, GAP, SIGNAL-TRANSDUCTION
512	1tx4	A	19	234	1.3e-39	0.44	1.00		P50-RHOGAP; CHAIN: A; TRANSFORMING PROTEIN RHOA; CHAIN: B;	COMPLEX(GTPASE ACTIVATN/PROTO-ONCOGENE) GTPASE-ACTIVATING PROTEIN RHOGAP; COMPLEX (GTPASE ACTIVATION/PROTO-ONCOGENE), GTPASE, 2 TRANSITION STATE, GAP
512	1tx4	A	19	234	1.4e-52			111.48	P50-RHOGAP; CHAIN: A; TRANSFORMING PROTEIN RHOA; CHAIN: B;	COMPLEX(GTPASE ACTIVATN/PROTO-ONCOGENE) GTPASE-ACTIVATING PROTEIN RHOGAP; COMPLEX (GTPASE ACTIVATION/PROTO-ONCOGENE), GTPASE, 2 TRANSITION STATE, GAP
512	1tx4	A	21	234	1.4e-52	0.67	1.00		P50-RHOGAP; CHAIN: A; TRANSFORMING PROTEIN RHOA; CHAIN: B;	COMPLEX(GTPASE ACTIVATN/PROTO-ONCOGENE) GTPASE-ACTIVATING PROTEIN RHOGAP; COMPLEX (GTPASE ACTIVATION/PROTO-ONCOGENE), GTPASE, 2 TRANSITION STATE, GAP
513	1d0s	A	322	635	3.6e-16	0.43	-0.19		NICOTINATE MONONUCLEOTIDE:5,6- CHAIN: A;	TRANSFERASE DINUCLEOTIDE-BINDING MOTIF, PHOSPHORIBOSYL TRANSFERASE
513	1kap	P	132	497	9e-11	0.81	-0.09		ALKALINE PROTEASE; 1KAP 4 CHAIN: P; 1KAP 5 TETRAPEPTIDE (GLY SER ASN SER); 1KAP 9 CHAIN: I; 1KAP 10	ZINC METALLOPROTEASE P. AERUGINOSA ALKALINE PROTEASE; 1KAP 6 CALCIUM BINDING PROTEIN 1KAP 19
513	1kap	P	225	691	7.2e-17			86.44	ALKALINE PROTEASE; 1KAP 4 CHAIN: P; 1KAP 5 TETRAPEPTIDE (GLY SER ASN SER); 1KAP 9 CHAIN: I; 1KAP 10	ZINC METALLOPROTEASE P. AERUGINOSA ALKALINE PROTEASE; 1KAP 6 CALCIUM BINDING PROTEIN 1KAP 19
513	1kap	P	237	654	7.2e-17	0.77	-0.18		ALKALINE PROTEASE; 1KAP 4 CHAIN: P; 1KAP 5 TETRAPEPTIDE (GLY SER ASN SER); 1KAP 9 CHAIN: I; 1KAP 10	ZINC METALLOPROTEASE P. AERUGINOSA ALKALINE PROTEASE; 1KAP 6 CALCIUM BINDING PROTEIN 1KAP 19
513	1kap	P	30	391	9e-15	0.77	-0.18		ALKALINE PROTEASE; 1KAP 4 CHAIN: P; 1KAP 5 TETRAPEPTIDE (GLY SER ASN SER); 1KAP 9	ZINC METALLOPROTEASE P. AERUGINOSA ALKALINE PROTEASE; 1KAP 6 CALCIUM BINDING PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: I; 1KAP 10	1KAP 19
513	1osm	A	9	322	7.2e-21	0.50	-0.20		OMPK36; CHAIN: A, B, C;	OUTER MEMBRANE PROTEIN OSMOPORIN; OUTER MEMBRANE PROTEIN, NON-SPECIFIC PORIN, OSMOPORIN, 2 BETA-BARREL, TRANSMEMBRANE
513	1pho		188	555	1.1e-21	0.81	-0.20		OUTER MEMBRANE PROTEIN PHOSPHOPORIN (PHOE) 1PHO 3	
517	1a1t	A	103	152	8e-11	-0.04	0.19		NUCLEOCAPSID PROTEIN; CHAIN: A; SL3 STEM-LOOP RNA; CHAIN: B;	COMPLEX (NUCLEOCAPSID PROTEIN/RNA) NUCLEOCAPSID PROTEIN, COMPLEX (NUCLEOCAPSID PROTEIN/RNA), 2 STEM-LOOP RNA
517	1a1t	A	124	182	1.3e-18	0.31	0.48		NUCLEOCAPSID PROTEIN; CHAIN: A; SL3 STEM-LOOP RNA; CHAIN: B;	COMPLEX (NUCLEOCAPSID PROTEIN/RNA) NUCLEOCAPSID PROTEIN, COMPLEX (NUCLEOCAPSID PROTEIN/RNA), 2 STEM-LOOP RNA
517	1a1t	A	50	99	8e-12	0.11	-0.06		NUCLEOCAPSID PROTEIN; CHAIN: A; SL3 STEM-LOOP RNA; CHAIN: B;	COMPLEX (NUCLEOCAPSID PROTEIN/RNA) NUCLEOCAPSID PROTEIN, COMPLEX (NUCLEOCAPSID PROTEIN/RNA), 2 STEM-LOOP RNA
517	1a1t	A	66	122	4.8e-13	0.11	0.01		NUCLEOCAPSID PROTEIN; CHAIN: A; SL3 STEM-LOOP RNA; CHAIN: B;	COMPLEX (NUCLEOCAPSID PROTEIN/RNA) NUCLEOCAPSID PROTEIN, COMPLEX (NUCLEOCAPSID PROTEIN/RNA), 2 STEM-LOOP RNA
517	1aaf		124	182	1.6e-18	0.29	-0.07		NUCLEOCAPSID PROTEIN HIV-1 NUCLEOCAPSID PROTEIN (MN ISOLATE) (NMR, 20 STRUCTURES) 1AAF 3	
517	1bj6	A	104	152	1.3e-10	0.16	0.18		DNA (ACGCC); CHAIN: D; NUCLEOCAPSID PROTEIN 7; CHAIN: A;	COMPLEX (NUCLEOCAPSID PROTEIN/DNA) (12-53)NCP7; COMPLEX (NUCLEOCAPSID PROTEIN/DNA), NUCLEIC ACID, 2 RETROVIRUS, VIRUS MORPHOGENESIS, ZINC FINGER
517	1bj6	A	134	180	9.6e-17	0.29	0.81		DNA (ACGCC); CHAIN: D; NUCLEOCAPSID PROTEIN 7; CHAIN: A;	COMPLEX (NUCLEOCAPSID PROTEIN/DNA) (12-53)NCP7; COMPLEX (NUCLEOCAPSID PROTEIN/DNA), NUCLEIC ACID, 2 RETROVIRUS, VIRUS

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										MORPHOGENESIS, ZINC FINGER
517	1bj6	A	58	97	1.6e-10	-0.20	0.03		DNA (ACGCC); CHAIN: D; NUCLEOCAPSID PROTEIN 7; CHAIN: A;	COMPLEX (NUCLEOCAPSID PROTEIN/DNA) (12-53)NCP7; COMPLEX (NUCLEOCAPSID PROTEIN/DNA), NUCLEIC ACID, 2 RETROVIRUS, VIRUS MORPHOGENESIS, ZINC FINGER
517	1bj6	A	77	122	9.6e-13	0.09	0.28		DNA (ACGCC); CHAIN: D; NUCLEOCAPSID PROTEIN 7; CHAIN: A;	COMPLEX (NUCLEOCAPSID PROTEIN/DNA) (12-53)NCP7; COMPLEX (NUCLEOCAPSID PROTEIN/DNA), NUCLEIC ACID, 2 RETROVIRUS, VIRUS MORPHOGENESIS, ZINC FINGER
517	1nc8		100	127	3.2e-05	-0.10	0.09		NUCLEOCAPSID PROTEIN; CHAIN: NULL;	NUCLEOCAPSID PROTEIN NUCLEOCAPSID PROTEIN, HIV-2, RNA RECOGNITION, ZINC FINGER
517	1nc8		129	156	1.6e-06	-0.29	0.04		NUCLEOCAPSID PROTEIN; CHAIN: NULL;	NUCLEOCAPSID PROTEIN NUCLEOCAPSID PROTEIN, HIV-2, RNA RECOGNITION, ZINC FINGER
517	1nc8		73	100	6.4e-06	-0.06	0.12		NUCLEOCAPSID PROTEIN; CHAIN: NULL;	NUCLEOCAPSID PROTEIN NUCLEOCAPSID PROTEIN, HIV-2, RNA RECOGNITION, ZINC FINGER
519	1mey	G	219	249	0.0056	-0.23	0.34		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
523	1axi	A	29	216	3.2e-52	0.59	1.00		GROWTH HORMONE; CHAIN: A; GROWTH HORMONE RECEPTOR; CHAIN: B;	COMPLEX (HORMONE/RECEPTOR) HGH; HGHBP; COMPLEX (HORMONE/RECEPTOR)
523	1axi	A	29	217	3.2e-52			243.17	GROWTH HORMONE; CHAIN: A; GROWTH HORMONE RECEPTOR; CHAIN: B;	COMPLEX (HORMONE/RECEPTOR) HGH; HGHBP; COMPLEX (HORMONE/RECEPTOR)
523	1bp3	A	27	216	1.6e-61	0.28	1.00		GROWTH HORMONE; CHAIN: A; PROLACTIN RECEPTOR; CHAIN: B;	HORMONE/GROWTH FACTOR HORMONE, RECEPTOR, HORMONE/GROWTH FACTOR
523	1bp3	A	27	216	1.6e-61			271.68	GROWTH HORMONE; CHAIN: A; PROLACTIN RECEPTOR; CHAIN:	HORMONE/GROWTH FACTOR HORMONE, RECEPTOR, HORMONE/GROWTH FACTOR

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									B;	
523	1hgu		28	216	1.6e-60	0.14	1.00		HUMAN GROWTH HORMONE; 1HGU 5 CHAIN: NULL; 1HGU 6	HORMONE HUMAN SOMATOTROPIN 1HGU 7 HORMONE 1HGU 11
523	1hgu		28	216	1.6e-60			264.37	HUMAN GROWTH HORMONE; 1HGU 5 CHAIN: NULL; 1HGU 6	HORMONE HUMAN SOMATOTROPIN 1HGU 7 HORMONE 1HGU 11
523	1hwg	A	27	216	4.8e-62	0.42	1.00		GROWTH HORMONE; CHAIN: A; GROWTH HORMONE BINDING PROTEIN; CHAIN: B, C;	COMPLEX (HORMONE/RECEPTOR) CYTOKINE, HORMONE, RECEPTOR, HEMATOPOIETIC, 2 COMPLEX (HORMONE/RECEPTOR)
523	1hwg	A	27	216	4.8e-62			272.96	GROWTH HORMONE; CHAIN: A; GROWTH HORMONE BINDING PROTEIN; CHAIN: B, C;	COMPLEX (HORMONE/RECEPTOR) CYTOKINE, HORMONE, RECEPTOR, HEMATOPOIETIC, 2 COMPLEX (HORMONE/RECEPTOR)
525	1ao7	D	15	100	0.0083	0.26	0.24		HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA-A2 HEAVY CHAIN; CLASS I MHC, T-CELL RECEPTOR, VIRAL PEPTIDE, 2 COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR
525	1hng	A	24	115	0.00018	0.25	0.23		T LYMPHOCYTE ADHESION GLYCOPROTEIN CD2 (RAT) 1HNG 3	
525	1qrn	D	15	100	0.0083	0.19	0.25		MHC CLASS I HLA-A; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE P6A; CHAIN: C; HMAN T-CELL RECEPTOR; CHAIN: D; HLA-A 0201; CHAIN: E;	IMMUNE SYSTEM HUMAN TCR/PEPTIDE/MHC COMPLEX, HLA-A2, HTLV-1, TAX, TCR, T 2 CELL RECEPTOR, IMMUNE SYSTEM
526	1one	A	2	431	0			509.15	ENOLASE; CHAIN: A, B;	LYASE 2-PHOSPHO-D-GLYCERATE HYDROLASE; LYASE, GLYCOLYSIS
526	1one	A	5	429	0	0.90	1.00		ENOLASE; CHAIN: A, B;	LYASE 2-PHOSPHO-D-GLYCERATE HYDROLASE; LYASE, GLYCOLYSIS
526	1pdz		2	431	0	1.07	1.00		ENOLASE; 1PDZ 4 CHAIN: NULL; 1PDZ 5	LYASE (CARBON-OXYGEN) 2-PHOSPHO-D-GLYCERATE DEHYDRATASE; 1PDZ 6
526	1pdz		2	432	0			563.45	ENOLASE; 1PDZ 4 CHAIN: NULL; 1PDZ 5	LYASE (CARBON-OXYGEN) 2-PHOSPHO-D-GLYCERATE DEHYDRATASE; 1PDZ 6

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
529	1b0x	A	14	62	5.4e-05	0.05	0.46		EPHA4 RECEPTOR TYROSINE KINASE; CHAIN: A;	TRANSFERASE RECEPTOR TYROSINE KINASE, PROTEIN INTERACTION MODULE, 2 DIMERIZATION DOMAIN, TRANSFERASE
529	1b4f	A	4	62	1.8e-06	0.74	1.00		EPHB2; CHAIN: A, B, C, D, E, F, G, H;	SIGNAL TRANSDUCTION SAM DOMAIN, EPH RECEPTOR, SIGNAL TRANSDUCTION, OLIGOMER
529	1sgg		3	62	9e-06	0.72	0.95		EPHRIN TYPE-B RECEPTOR 2; CHAIN: NULL;	TYROSINE-PROTEIN KINASE NMR, RECEPTOR OLIGOMERIZATION, EPH RECEPTORS, TYROSINE 2 PHOSPHORYLATION, SIGNAL TRANSDUCTION, TYROSINE-PROTEIN 3 KINASE
531	1b6t	A	7	157	8e-33	-0.13	0.36		PHOSPHOPANTETHEINE ADENYLYLTRANSFERASE; CHAIN: A, B;	TRANSFERASE PPAT, KDTB; COENZYME A BIOSYNTHESIS
536	1d0s	A	166	525	7.2e-13	0.52	-0.20		NICOTINATE MONONUCLEOTIDE:5,6- CHAIN: A;	TRANSFERASE DINUCLEOTIDE-BINDING MOTIF, PHOSPHORIBOSYL TRANSFERASE
536	1osm	A	27	392	7.2e-20	0.38	-0.20		OMPK36; CHAIN: A, B, C;	OUTER MEMBRANE PROTEIN OSMOPORIN; OUTER MEMBRANE PROTEIN, NON-SPECIFIC PORIN, OSMOPORIN, 2 BETA-BARREL, TRANSMEMBRANE
536	1osm	A	301	631	1.3e-19	0.74	-0.20		OMPK36; CHAIN: A, B, C;	OUTER MEMBRANE PROTEIN OSMOPORIN; OUTER MEMBRANE PROTEIN, NON-SPECIFIC PORIN, OSMOPORIN, 2 BETA-BARREL, TRANSMEMBRANE
536	1pho		316	629	5.4e-15	0.83	-0.20		OUTER MEMBRANE PROTEIN PHOSPHOPORIN (PHOE) 1PHO 3	
536	2omf		282	629	9e-12	0.82	-0.20		MATRIX PORIN OUTER MEMBRANE PROTEIN F; 2OMF 5 CHAIN: NULL; 2OMF 6	INTEGRAL MEMBRANE PROTEIN PORIN MATRIX PORIN, OMPF PORIN; 2OMF 7 PORIN, MEMBRANE PROTEIN 2OMF 12

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
537	1a26		51	160	7.2e-06	-0.24	0.05		POLY (ADP-RIBOSE) POLYMERASE; CHAIN: NULL;	TRANSFERASE PARP-CF, POLY(ADP-RIBOSE) TRANSFERASE, POLY TRANSFERASE, GLYCOSYLTRANSFERASE, NAD(+) 2 ADP-RIBOSYLTRANSFERASE
538	1a26		51	160	7.2e-06	-0.24	0.05		POLY (ADP-RIBOSE) POLYMERASE; CHAIN: NULL;	TRANSFERASE PARP-CF, POLY(ADP-RIBOSE) TRANSFERASE, POLY TRANSFERASE, GLYCOSYLTRANSFERASE, NAD(+) 2 ADP-RIBOSYLTRANSFERASE
540	1fao	A	1	91	5.4e-19	-0.06	0.27		DUAL ADAPTOR OF PHOSPHOTYROSINE AND 3-CHAIN: A;	SIGNALING PROTEIN DAPP1, PHISH, BAM32; PLECKSTRIN, 3-PHOSPHOINOSITIDES, INOSITOL TETRAKISPHOSPHATE 2 SIGNAL TRANSDUCTION PROTEIN, ADAPTOR PROTEIN
540	1fao	A	2	86	4.8e-11	0.51	0.33		DUAL ADAPTOR OF PHOSPHOTYROSINE AND 3-CHAIN: A;	SIGNALING PROTEIN DAPP1, PHISH, BAM32; PLECKSTRIN, 3-PHOSPHOINOSITIDES, INOSITOL TETRAKISPHOSPHATE 2 SIGNAL TRANSDUCTION PROTEIN, ADAPTOR PROTEIN
540	1fb8	A	1	91	1.8e-19	0.35	0.46		DUAL ADAPTOR OF PHOSPHOTYROSINE AND 3-CHAIN: A;	SIGNALING PROTEIN DAPP1, PHISH, BAM32; PLECKSTRIN, 3-PHOSPHOINOSITIDES, INOSITOL TETRAKISPHOSPHATE 2 SIGNAL TRANSDUCTION PROTEIN, ADAPTOR PROTEIN
540	1fb8	A	2	86	4.8e-11	0.00	0.13		DUAL ADAPTOR OF PHOSPHOTYROSINE AND 3-CHAIN: A;	SIGNALING PROTEIN DAPP1, PHISH, BAM32; PLECKSTRIN, 3-PHOSPHOINOSITIDES, INOSITOL TETRAKISPHOSPHATE 2 SIGNAL TRANSDUCTION PROTEIN, ADAPTOR PROTEIN
540	1fgy	A	2	91	1.6e-15	0.67	0.54		GRP1; CHAIN: A;	SIGNALING PROTEIN ARF1 GUANINE NUCLEOTIDE EXCHANGE FACTOR AND

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										PH DOMAIN
540	1fgy	A	2	96	9.6e-15	0.57	0.07		GRP1; CHAIN: A;	SIGNALING PROTEIN ARF1 GUANINE NUCLEOTIDE EXCHANGE FACTOR AND PH DOMAIN
540	1pls		1	95	1.3e-17	0.74	0.21		PHOSPHORYLATION PLECKSTRIN (N-TERMINAL PLECKSTRIN HOMOLOGY DOMAIN) MUTANT 1PLS 3 WITH LEU GLU (HIS)6 ADDED TO THE C TERMINUS 1PLS 4 (INS(G105-LEHHHHHH)) (NMR, 25 STRUCTURES) 1PLS 5	
540	1pls		2	96	6.4e-11	0.30	0.07		PHOSPHORYLATION PLECKSTRIN (N-TERMINAL PLECKSTRIN HOMOLOGY DOMAIN) MUTANT 1PLS 3 WITH LEU GLU (HIS)6 ADDED TO THE C TERMINUS 1PLS 4 (INS(G105-LEHHHHHH)) (NMR, 25 STRUCTURES) 1PLS 5	
540	1pms		1	88	1.8e-14	0.89	0.25		SOS 1; CHAIN: NULL;	SIGNAL TRANSDUCTION SON OF SEVENLESS; PLECKSTRIN, SON OF SEVENLESS, SIGNAL TRANSDUCTION
541	1et7	A	35	137	0.0051	-0.04	0.01		NITRITE REDUCTASE; CHAIN: A;	OXIDOREDUCTASE CU-NIR; GREEK KEY BETA BARREL DOMAIN
542	1ae1	A	3	245	3.6e-67			117.73	TROPINONE REDUCTASE-I; CHAIN: A, B;	OXIDOREDUCTASE OXIDOREDUCTASE, TROPANE ALKALOID BIOSYNTHESIS, REDUCTION OF 2 TROPINONE TO TROPINE, SHORT-CHAIN DEHYDROGENASE
542	1ae1	A	4	245	3.6e-67	0.61	1.00		TROPINONE REDUCTASE-I; CHAIN: A, B;	OXIDOREDUCTASE OXIDOREDUCTASE, TROPANE ALKALOID BIOSYNTHESIS, REDUCTION OF 2 TROPINONE TO TROPINE, SHORT-CHAIN DEHYDROGENASE
542	1ae1	B	3	245	1.1e-71			134.35	TROPINONE REDUCTASE-I;	OXIDOREDUCTASE OXIDOREDUCTASE,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: A, B;	TROPANE ALKALOID BIOSYNTHESIS, REDUCTION OF 2 TROPINONE TO TROPINE, SHORT-CHAIN DEHYDROGENASE
542	1ael	B	4	244	1.1e-65	0.37	1.00		TROPINONE REDUCTASE-I; CHAIN: A, B;	OXIDOREDUCTASE OXIDOREDUCTASE, TROPANE ALKALOID BIOSYNTHESIS, REDUCTION OF 2 TROPINONE TO TROPINE, SHORT-CHAIN DEHYDROGENASE
542	1ael	B	4	245	1.1e-71	0.69	1.00		TROPINONE REDUCTASE-I; CHAIN: A, B;	OXIDOREDUCTASE OXIDOREDUCTASE, TROPANE ALKALOID BIOSYNTHESIS, REDUCTION OF 2 TROPINONE TO TROPINE, SHORT-CHAIN DEHYDROGENASE
542	1bdb		3	245	7.2e-65	0.57	1.00		CIS-BIPHENYL-2,3-DIHYDRODIOL-2,3-DEHYDROGENASE; CHAIN: NULL;	OXIDOREDUCTASE NAD-DEPENDENT OXIDOREDUCTASE, SHORT-CHAIN ALCOHOL 2 DEHYDROGENASE, PCB DEGRADATION
542	1bdb		3	245	7.2e-65			82.09	CIS-BIPHENYL-2,3-DIHYDRODIOL-2,3-DEHYDROGENASE; CHAIN: NULL;	OXIDOREDUCTASE NAD-DEPENDENT OXIDOREDUCTASE, SHORT-CHAIN ALCOHOL 2 DEHYDROGENASE, PCB DEGRADATION
542	1bxk	A	7	150	7.2e-05	0.14	0.41		DTDP-GLUCOSE 4,6-DEHYDRATASE; CHAIN: A, B;	LYASE EPIMERASE, DEHYDRATASE, DEHYDROGENASE, LYASE
542	1cyd	A	3	243	5.4e-71	0.67	1.00		CARBONYL REDUCTASE; CHAIN: A, B, C, D;	OXIDOREDUCTASE SHORT-CHAIN DEHYDROGENASE, OXIDOREDUCTASE
542	1cyd	A	3	245	5.4e-71			117.07	CARBONYL REDUCTASE; CHAIN: A, B, C, D;	OXIDOREDUCTASE SHORT-CHAIN DEHYDROGENASE, OXIDOREDUCTASE
542	1db3	A	7	180	0.0018	0.07	0.93		GDP-MANNOSE 4,6-DEHYDRATASE; CHAIN: A;	LYASE DEHYDRATASE, NADP, GDP-MANNOSE, GDP-FUCOSE
542	1ek6	A	7	135	0.00036	0.05	0.69		UDP-GALACTOSE 4-EPIMERASE; CHAIN: A, B;	ISOMERASE EPIMERASE, SHORT-CHAIN DEHYDROGENASE, GALACTOSEMIA
542	1eny		1	242	3.6e-61			66.42	ENOYL-ACYL CARRIER PROTEIN (ACP) REDUCTASE; 1ENY 4 CHAIN: NULL; 1ENY 5	OXIDOREDUCTASE INHA; 1ENY 6
542	1fds		5	243	3.2e-28			57.21	17-BETA-HYDROXYSTEROID-DEHYDROGENASE; CHAIN: NULL;	DEHYDROGENASE DEHYDROGENASE, 17-BETA-HYDROXYSTEROID

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
542	1fmc	A	3	242	1.1e-63	0.72	1.00		7 ALPHA-HYDROXYSTEROID DEHYDROGENASE; CHAIN: A, B;	OXIDOREDUCTASE SHORT-CHAIN DEHYDROGENASE/REDUCTASE, BILE ACID CATABOLISM
542	1fmc	A	3	245	1.1e-63			113.82	7 ALPHA-HYDROXYSTEROID DEHYDROGENASE; CHAIN: A, B;	OXIDOREDUCTASE SHORT-CHAIN DEHYDROGENASE/REDUCTASE, BILE ACID CATABOLISM
542	1gdh	A	4	58	0.0014	0.12	0.51		OXIDOREDUCTASE(CHOH (D)-NAD(P)+ (A)) D-GLYCERATE DEHYDROGENASE (APO FORM) (E.C.1.1.1.29) 1GDH 3	
542	1hdc	A	2	244	3.2e-66	0.45	1.00		OXIDOREDUCTASE 3-ALPHA, 20-BETA-HYDROXYSTEROID DEHYDROGENASE (E.C.1.1.1.53) 1HDC 3 COMPLEXED WITH CARBENOXOLONE 1HDC 4	
542	1hdc	A	2	245	3.2e-66			110.09	OXIDOREDUCTASE 3-ALPHA, 20-BETA-HYDROXYSTEROID DEHYDROGENASE (E.C.1.1.1.53) 1HDC 3 COMPLEXED WITH CARBENOXOLONE 1HDC 4	
542	1oaa		1	241	1.8e-57			57.85	SEPIAPTERIN REDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE SEPIAPTERIN REDUCTASE, TETRAHYDROBIOPTERIN, OXIDOREDUCTASE
542	1qor	A	6	83	3.6e-07	0.55	0.99		OXIDOREDUCTASE QUINONE OXIDOREDUCTASE COMPLEXED WITH NADPH 1QOR 3	
542	1qrr	A	7	154	0.00054	-0.26	0.27		SULFOLIPID BIOSYNTHESIS (SQD1) PROTEIN; CHAIN: A;	ISOMERASE ROSSMANN FOLD, SHORT HYDROGEN BONDS, SDR HOMOLOG, ISOMERASE
542	1qsg	A	2	245	1.3e-63	0.63	1.00		ENOYL-REDUCTASE; CHAIN: A, B, C, D, E, F, G, H;	OXIDOREDUCTASE ENOYL REDUCTASE, OXIDOREDUCTASE
542	1ybv	A	1	245	3.6e-66			106.51	TRIHYDROXYNAPHTHALENE REDUCTASE; CHAIN: A, B;	OXIDOREDUCTASE NAPHTHOL REDUCTASE; OXIDOREDUCTASE
542	1ybv	A	3	245	3.6e-66	0.72	1.00		TRIHYDROXYNAPHTHALENE REDUCTASE; CHAIN: A, B;	OXIDOREDUCTASE NAPHTHOL REDUCTASE; OXIDOREDUCTASE
542	2ae2	A	1	245	5.4e-71			114.97	TROPINONE REDUCTASE-II; CHAIN: A, B;	OXIDOREDUCTASE OXIDOREDUCTASE, TROPANE ALKALOID BIOSYNTHESIS, REDUCTION OF 2 TROPINONE TO

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										PSEUDOTROPINE, SHORT-CHAIN DEHYDROGENASE
542	2ae2	A	3	245	5.4e-71	0.68	1.00		TROPINONE REDUCTASE-II; CHAIN: A, B;	OXIDOREDUCTASE OXIDOREDUCTASE, TROPANE ALKALOID BIOSYNTHESIS, REDUCTION OF 2 TROPINONE TO PSEUDOTROPINE, SHORT-CHAIN DEHYDROGENASE
542	2pgd		18	48	0.0036	-0.61	0.12		OXIDOREDUCTASE (CHOH(D)-NADP+(A)) 6-PHOSPHOGLUCONATE DEHYDROGENASE (6-PGDH) (E.C.1.1.1.44) 2PGD 3	
549	1bs2	A	16	241	3.2e-75	0.06	1.00		ARGINYL-TRNA SYNTHETASE; CHAIN: A;	LIGASE ARGRS, ARGININE - TRNA LIGASE; LIGASE, AMINOACYL-TRNA SYNTHETASE, PROTEIN BIOSYNTHESIS
549	1bs2	A	1	403	0			257.02	ARGINYL-TRNA SYNTHETASE; CHAIN: A;	LIGASE ARGRS, ARGININE - TRNA LIGASE; LIGASE, AMINOACYL-TRNA SYNTHETASE, PROTEIN BIOSYNTHESIS
549	1bs2	A	16	403	0	0.38	1.00		ARGINYL-TRNA SYNTHETASE; CHAIN: A;	LIGASE ARGRS, ARGININE - TRNA LIGASE; LIGASE, AMINOACYL-TRNA SYNTHETASE, PROTEIN BIOSYNTHESIS
549	1qu2	A	40	114	1.8e-07	-0.51	0.00		ISOLEUCYL-TRNA SYNTHETASE; CHAIN: A; ISOLEUCYL-TRNA; CHAIN: T;	LIGASE/RNA ISOLEUCINE--TRNA LIGASE, ILERS; PROTEIN-RNA COMPLEX, METAL IONS, EDITING TRNA SYNTHETASE, 2 DOUBLE-SIEVE
550	1bs2	A	16	241	3.2e-75	0.06	1.00		ARGINYL-TRNA SYNTHETASE; CHAIN: A;	LIGASE ARGRS, ARGININE - TRNA LIGASE; LIGASE, AMINOACYL-TRNA SYNTHETASE, PROTEIN BIOSYNTHESIS
550	1bs2	A	1	403	0			257.02	ARGINYL-TRNA SYNTHETASE; CHAIN: A;	LIGASE ARGRS, ARGININE - TRNA LIGASE; LIGASE, AMINOACYL-TRNA SYNTHETASE, PROTEIN BIOSYNTHESIS
550	1bs2	A	16	403	0	0.38	1.00		ARGINYL-TRNA SYNTHETASE; CHAIN: A;	LIGASE ARGRS, ARGININE - TRNA LIGASE; LIGASE, AMINOACYL-TRNA SYNTHETASE, PROTEIN BIOSYNTHESIS
550	1qu2	A	40	114	1.8e-07	-0.51	0.00		ISOLEUCYL-TRNA SYNTHETASE; CHAIN: A; ISOLEUCYL-TRNA;	LIGASE/RNA ISOLEUCINE--TRNA LIGASE, ILERS; PROTEIN-RNA COMPLEX, METAL

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: T;	IONS, EDITING TRNA SYNTHETASE, 2 DOUBLE-SIEVE
552	1ek9	A	95	255	9e-06	-0.22	0.22		OUTER MEMBRANE PROTEIN TOLC; CHAIN: A, B, C;	MEMBRANE PROTEIN INTEGRAL MEMBRANE PROTEIN, ALPHA HELICAL BARREL, BETA BARREL
552	1qu7	A	88	146	3.6e-13	0.39	-0.20		METHYL-ACCEPTING CHEMOTAXIS PROTEIN I; CHAIN: A, B;	SIGNALING PROTEIN SERINE, CHEMOTAXIS, FOUR HELICAL-BUNDLE
555	1dvp	A	401	457	7.2e-06	-0.87	0.46		HEPATOCYTE GROWTH FACTOR-REGULATED TYROSINE CHAIN: A;	TRANSFERASE HRS; HRS, VHS, FYVE, ZINC FINGER, SUPERHELIX
555	1ptq		421	452	0.0072	-0.82	0.05		PROTEIN KINASE C DELTA TYPE; 1PTQ 4	PHOSPHOTRANSFERASE
555	1tbn		421	453	0.0054	-0.83	0.07		PROTEIN KINASE C, GAMMA TYPE; CHAIN: NULL;	CALCIUM-BINDING PROTEIN RAT BRAIN PKC-G; CALCIUM-BINDING PROTEIN, PROTEIN KINASE C, PKC, TRANSFERASE
555	1vfy	A	421	457	5.4e-05	-0.84	0.05		PHOSPHATIDYLINOSITOL-3-PHOSPHATE BINDING FYVE CHAIN: A;	TRANSPORT PROTEIN FYVE DOMAIN, ENDOSOME MATURATION, INTRACELLULAR TRAFFICKING, 2 TRANSPORT PROTEIN
555	1zbd	B	385	452	9e-05	-0.92	0.22		RAB-3A; CHAIN: A; RABPHILIN-3A; CHAIN: B;	COMPLEX (GTP-BINDING/EFFECTOR) RAS-RELATED PROTEIN RAB3A; COMPLEX (GTP-BINDING/EFFECTOR), G PROTEIN, EFFECTOR, RABCDR, 2 SYNAPTIC EXOCYTOSIS, RAB PROTEIN, RAB3A, RABPHILIN
556	1chc		122	179	3.6e-05	-0.16	0.01		VIRUS EQUINE HERPES VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	
564	1cwv	A	225	337	0.0035	-0.87	0.06		INVASIN; CHAIN: A;	STRUCTURAL PROTEIN INTEGRIN-BINDING PROTEIN, INV GENE
564	1dan	L	3758	3833	3.2e-13	0.02	-0.17		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
564	1edh	A	1029	1237	1.8e-51			114.54	E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	1043	1237	1.8e-51	0.57	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	1068	1237	4.8e-36	0.49	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	1147	1343	3.6e-38	0.33	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	1149	1347	3.2e-31	0.37	0.77		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	1287	1451	8e-24	0.41	0.49		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	1367	1554	5.4e-34	0.34	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	1384	1554	4.8e-33	0.14	0.93		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	1459	1662	1.6e-57	0.29	0.98		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										CALCIUM BINDING PROTEIN
564	1edh	A	1578	1760	3.6e-35	0.50	0.87		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	1596	1760	6.4e-33	0.59	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	1676	1874	3.6e-38	0.37	0.54		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	1700	1874	8e-32	0.09	0.99		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	1789	1972	9e-31	0.19	0.52		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	183	359	6.4e-28	-0.06	0.04		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	1896	2076	7.2e-25	0.46	0.89		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	1915	2076	1.6e-20	0.21	0.43		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	1991	2177	1.8e-28	0.03	0.81		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	2015	2177	1.4e-19	-0.24	0.69		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	2094	2261	4.8e-22	-0.34	0.47		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	2100	2278	7.2e-25	0.30	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	2188	2367	4.8e-38	0.01	0.89		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	2191	2385	3.6e-40	0.29	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	2302	2485	1.6e-34	0.42	0.68		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	2319	2487	6.4e-32	0.33	0.59		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	2427	2591	4.8e-34	0.17	0.46		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	2529	2697	8e-30	0.08	0.89		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	256	454	1.1e-14	0.23	0.80		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										CALCIUM BINDING PROTEIN
564	1edh	A	2609	2803	3.2e-41	0.35	0.43		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	2714	2912	3.6e-38	0.30	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	2738	2912	9.6e-38	0.45	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	2818	3017	1.4e-31	0.34	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	2819	3017	7.2e-38	0.50	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	292	455	3.2e-25	0.31	0.82		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	2929	3119	3.6e-37	0.28	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	2955	3119	3.2e-28	0.13	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	3031	3224	4.8e-46	0.50	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	3134	3329	3.6e-47	0.33	0.99		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	3137	3318	3.2e-26	0.04	0.88		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	3248	3434	7.2e-47	0.63	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	3266	3434	3.2e-18	0.26	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	3345	3537	1.8e-38	0.50	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	3372	3539	6.4e-17	-0.00	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	3448	3634	1.1e-18	-0.18	0.34		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	375	559	1.1e-27	0.38	0.98		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	40	249	6.4e-50	-0.04	0.70		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	406	561	3.2e-30	0.01	0.29		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										CALCIUM BINDING PROTEIN
564	1edh	A	474	663	1.6e-26	0.25	0.76		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	498	667	1.1e-25	0.12	0.42		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	597	814	6.4e-20	-0.09	0.28		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	720	919	8e-51	0.41	0.95		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	829	1024	7.2e-43	0.46	1.00		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	856	1016	6.4e-31	0.16	0.49		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	928	1131	3.6e-40	0.29	0.88		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1edh	A	964	1131	3.2e-27	0.41	0.72		E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
564	1emn		3750	3830	3.2e-14	0.02	-0.14		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
564	1fak	L	3758	3833	3.2e-13	0.00	-0.14		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
564	1ncg		1043	1129	9e-19	0.64	0.36		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		1066	1129	1.1e-06	0.35	0.05		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		1142	1236	3.6e-18	0.51	0.81		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		1146	1235	0.00048	0.51	0.94		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		1247	1344	7.2e-12	0.58	0.83		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		1359	1447	9e-11	0.61	0.25		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		1457	1553	1.1e-19	0.22	0.88		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		153	247	5.4e-15	0.37	0.37		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		156	235	1.6e-05	0.28	0.52		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		1565	1660	7.2e-19	0.54	0.36		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		1594	1660	8e-07	0.24	0.05		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		1675	1757	3.6e-11	0.29	-0.09		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		1680	1758	0.00014	0.43	0.05		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		1770	1873	9e-16	-0.00	0.28		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
564	1ncg		1807	1873	3.2e-06	-0.23	0.16		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		1984	2074	1.3e-09	0.13	0.69		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		2088	2178	1.8e-10	0.38	-0.07		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		2185	2270	1.4e-14	-0.28	0.18		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		2288	2383	1.8e-18	0.11	0.34		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		2394	2486	1.4e-08	0.23	0.06		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		2497	2590	9e-11	0.02	0.51		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		2604	2696	1.3e-11	0.38	0.23		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		2707	2801	3.2e-09	0.42	0.12		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		2708	2801	9e-12	0.40	0.30		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		2812	2910	5.4e-20	0.26	0.18		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		2818	2897	1.6e-05	0.19	0.22		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		2924	3016	9e-14	0.30	0.48		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		3026	3118	1.4e-18	0.25	0.75		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		3026	3120	3.2e-15	0.40	0.59		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		3131	3223	1.1e-18	0.34	0.68		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		3137	3209	1.6e-05	0:06	0.42		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		3237	3327	3.6e-19	0.31	0.04		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		3264	3311	0.00064	0.19	0.01		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
564	1ncg		3339	3432	1.3e-19	0.22	0.48		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		3447	3538	1.1e-10	-0.09	0.41		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		3559	3634	0.00018	0.46	0.11		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		465	557	9e-14	0.22	0.09		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		583	663	3.6e-07	0.16	0.43		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		597	666	0.0046	-0.05	0.12		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		718	813	1.3e-17	0.06	0.55		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		824	915	1.6e-17	0.10	0.29		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		932	1023	1.8e-20	0.53	0.42		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1ncg		936	1009	3.2e-05	0.36	0.46		N-CADHERIN; 1NCG 3	CELL ADHESION PROTEIN CADHERIN 1NCG 13
564	1nci	B	1043	1131	1.3e-17	0.41	0.51		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	1067	1131	1.3e-07	0.34	0.65		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	1133	1237	1.6e-17	0.42	0.72		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	1174	1237	0.0016	0.20	0.69		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	1239	1344	1.8e-07	0.10	0.64		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	1359	1448	1.8e-11	0.63	0.93		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	1387	1449	3.2e-05	0.39	0.03		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	1457	1554	1.3e-17	-0.03	0.83		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	154	248	5.4e-13	0.07	0.23		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
564	1nci	B	1556	1660	5.4e-17	-0.06	0.28		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	1595	1662	1.6e-07	0.07	0.17		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	1675	1760	1.8e-14	0.55	0.33		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	1680	1763	3.2e-05	0.33	0.04		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	1762	1874	5.4e-13	0.01	0.09		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	1808	1874	8e-07	-0.41	0.11		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	182	249	1.6e-05	-0.32	0.19		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	1887	1972	1.1e-05	0.44	0.35		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	1986	2076	3.6e-10	0.31	0.93		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	2078	2178	3.6e-09	0.19	0.00		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	2185	2270	3.2e-12	-0.40	0.09		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	2186	2275	7.2e-15	0.06	0.39		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	2280	2385	7.2e-18	-0.03	0.78		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	2394	2486	3.6e-06	0.47	0.37		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	2500	2591	9e-10	0.40	0.96		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	2593	2696	3.6e-11	0.32	0.03		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	2707	2803	1.6e-09	0.56	0.95		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	2710	2803	1.4e-12	0.24	0.95		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	2805	2912	3.6e-19	0.19	0.65		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
564	1nci	B	2846	2912	9.6e-06	-0.16	0.16		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	2924	3017	1.8e-13	0.35	0.62		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	3019	3119	1.6e-17	0.19	0.77		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	3026	3120	8e-13	0.05	0.37		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	3121	3224	3.6e-17	0.56	0.55		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	3162	3224	6.4e-05	-0.20	0.39		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	3226	3329	1.8e-22	0.31	0.23		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	3331	3434	1.8e-19	0.23	0.72		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	3381	3434	0.0046	0.24	1.00		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	3436	3538	1.8e-08	-0.43	0.25		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	368	454	3.6e-10	0.43	0.95		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	40	141	8e-15	0.06	-0.14		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	563	663	9e-05	0.30	0.65		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	718	814	8e-16	0.05	0.86		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	816	919	7.2e-15	0.24	0.22		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	855	919	4.8e-06	0.44	0.37		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	921	1024	3.6e-19	0.38	0.57		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1nci	B	936	1009	0.00014	0.03	0.18		N-CADHERIN; 1NCI 3	CELL ADHESION PROTEIN CADHERIN 1NCI 13
564	1ncj	A	1062	1237	1.6e-38	0.77	0.99		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
564	1ncj	A	1146	1347	1.6e-32	0.59	1.00		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	1278	1448	4.8e-24	0.68	0.98		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	1355	1554	6.4e-33	-0.08	0.68		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	1458	1662	6.4e-62	0.30	0.94		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	156	359	1.1e-29	0.07	-0.06		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	1571	1760	6.4e-33	0.43	1.00		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	1680	1874	1.3e-34	0.38	0.98		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	1777	1977	9.6e-34	-0.07	0.25		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	1910	2079	6.4e-19	0.51	0.22		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	2008	2177	1.3e-22	-0.12	0.45		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	2091	2266	1.6e-22	-0.26	0.22		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	2185	2371	1.4e-42	-0.08	0.76		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	2293	2487	1.3e-32	0.33	0.57		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	2400	2591	1.3e-36	0.21	0.82		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	2504	2697	1.6e-32	0.27	0.25		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	2606	2803	1.6e-45	0.40	0.71		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	2707	2912	4.8e-43	0.39	1.00		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	2818	3017	3.2e-33	0.30	1.00		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	2927	3119	1.6e-28	0.61	1.00		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
564	1ncj	A	299	455	1.1e-25	0.02	0.34		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	3027	3224	9.6e-53	0.42	1.00		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	3028	3223	9.6e-53			110.81	N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	3153	3316	8e-30	0.38	1.00		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	3238	3434	1.6e-18	0.51	0.99		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	3365	3539	4.8e-19	0.39	1.00		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	374	561	9.6e-34	0.17	0.60		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	40	249	3.2e-56	0.02	0.37		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	489	667	1.4e-29	0.33	0.87		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	595	814	1.3e-22	-0.30	0.10		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	719	919	6.4e-56	0.09	1.00		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	829	1010	1.6e-33	0.11	0.40		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1ncj	A	936	1131	6.4e-31	0.39	0.87		N-CADHERIN; CHAIN: A;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN
564	1pfx	L	3794	3883	3.2e-12	0.16	-0.08		FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
564	1qfk	L	3762	3833	4.8e-12	0.10	-0.01		COAGULATION FACTOR VIIA (LIGHT CHAIN); CHAIN: L; COAGULATION FACTOR VIIA (HEAVY CHAIN); CHAIN: H; TRIPEPTIDYL INHIBITOR; CHAIN:	SERINE PROTEASE FVIIA; FVIIA; BLOOD COAGULATION, SERINE PROTEASE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									C;	
564	1qfk	L	3795	3883	6.4e-15	0.03	-0.19		COAGULATION FACTOR VIIA (LIGHT CHAIN); CHAIN: L; COAGULATION FACTOR VIIA (HEAVY CHAIN); CHAIN: H; TRIPEPTIDYL INHIBITOR; CHAIN: C;	SERINE PROTEASE FVIIA; FVIIA; BLOOD COAGULATION, SERINE PROTEASE
564	1suh		1043	1135	7.2e-22	0.47	0.53		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		1147	1241	1.8e-20	0.58	0.99		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		1149	1241	3.2e-07	0.31	0.82		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		1265	1343	1.3e-09	0.41	0.52		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		1363	1452	1.8e-12	0.65	0.88		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		1457	1558	1.3e-22	-0.04	0.65		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		156	251	9e-14	0.00	0.07		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		1578	1665	3.6e-18	0.27	0.24		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		1596	1666	1.3e-09	0.33	0.41		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		1676	1764	1.3e-13	0.58	0.09		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		1700	1764	1.6e-08	0.18	0.36		EPITHELIAL CADHERIN; CHAIN:	CELL ADHESION UVOMORULIN;

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									NULL;	CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		1790	1877	1.3e-17	-0.13	0.07		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		1896	1973	1.3e-07	0.33	0.03		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		1991	2080	1.6e-11	0.05	0.88		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		2100	2178	1.8e-10	0.02	-0.08		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		2185	2271	8e-14	-0.02	0.34		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		2193	2282	7.2e-17	0.17	0.95		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		2302	2389	7.2e-18	0.43	0.52		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		2397	2489	1.8e-05	0.42	0.21		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		2502	2595	1.4e-12	0.29	0.47		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		2529	2595	4.8e-07	-0.46	0.43		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		2609	2701	1.3e-09	0.37	-0.15		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		2622	2696	3.6e-10	0.31	-0.17		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
564	1suh		2711	2807	9e-13	0.34	0.88		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		2738	2807	4.8e-10	0.19	0.70		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		2812	2915	7.2e-22	0.23	0.81		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		292	363	0.0021	0.09	0.06		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		2929	3021	3.6e-15	0.18	0.98		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		2955	3021	0.00011	-0.04	0.90		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		3026	3123	3.2e-15	0.17	0.69		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		3034	3123	3.6e-20	0.60	1.00		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		3131	3228	1.4e-20	0.28	0.69		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		3137	3228	4.8e-06	0.05	0.22		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		3248	3333	7.2e-21	0.42	0.40		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		3345	3438	3.6e-21	0.50	0.75		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		3372	3438	0.0046	0.39	1.00		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										ADHESION
564	1suh		40	145	1.6e-18	-0.22	0.06		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		406	459	3.2e-06	-0.21	0.05		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		474	565	1.4e-14	-0.33	0.12		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		498	565	0.0016	-0.21	0.05		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		585	663	0.00011	0.25	0.60		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		718	818	3.2e-21	0.35	0.57		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		829	923	1.3e-17	0.26	0.62		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		856	923	1.6e-07	-0.10	0.24		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		934	1028	1.6e-19	0.23	0.29		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
564	1suh		964	999	0.0035	-0.51	0.11		EPITHELIAL CADHERIN; CHAIN: NULL;	CELL ADHESION UVOMORULIN; CADHERIN, CALCIUM BINDING, CELL ADHESION
565	1qf6	A	65	323	6.4e-67	0.24	1.00		THREONYL-TRNA SYNTHETASE; CHAIN: A; THREONINE TRNA; CHAIN: B;	LIGASE/RNA THRRS; TRNA (THR); THREONYL-TRNA SYNTHETASE, TRNA(THR), AMP, ZINC, MRNA, 2 AMINOACYLATION, TRANSLATIONAL REGULATION, PROTEIN/RNA

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
568	1a70		62	166	4.8e-27	-0.06	0.49		FERREDOXIN; CHAIN: NULL;	IRON-SULFUR PROTEIN IRON-SULFUR PROTEIN, PHOTOSYNTHESIS, ELECTRON TRANSPORT TER
568	1awd		64	166	3.2e-28	-0.09	0.66		FERREDOXIN; CHAIN: NULL	ELECTRON TRANSPORT ELECTRON TRANSPORT, EUKARYOTIC, GREEN ALGA, ELECTRON 2 TRANSFER, METALLOPROTEIN
568	1ayf	A	63	163	5.4e-31			87.51	ADRENODOXIN; CHAIN: A, B;	ELECTRON TRANSPORT [2FE-2S]FERREDOXIN, ADRENODOXIN, ELECTRON TRANSPORT
568	1ayf	A	64	163	5.4e-31	0.87	1.00		ADRENODOXIN; CHAIN: A, B;	ELECTRON TRANSPORT [2FE-2S]FERREDOXIN, ADRENODOXIN, ELECTRON TRANSPORT
568	1b9r	A	65	166	1.8e-28			66.15	TERPREDOXIN; CHAIN: A;	FERREDOXIN STRUCTURE FROM MOLMOL, FERREDOXIN
568	1b9r	A	66	163	1.8e-28	0.20	1.00		TERPREDOXIN; CHAIN: A;	FERREDOXIN STRUCTURE FROM MOLMOL, FERREDOXIN
568	1czp	A	62	166	4.8e-29	0.05	0.30		FERREDOXIN I; CHAIN: A, B	ELECTRON TRANSPORT [2FE-2S] PROTEIN, CRYSTAL REDUCED WITH DITHIONITE
568	1fxi	A	62	166	6.4e-29	-0.08	0.22		ELECTRON TRANSFER (IRON-SULFUR PROTEIN) FERREDOXIN I 1FXI 3	
568	1gpx		64	167	1.6e-20			82.36	PUTIDAREDOXIN; CHAIN: NULL;	ELECTRON TRANSPORT C85S GAPDX; ELECTRON TRANSPORT, GAPDX C85S, 20 STRUCTURES ALIGNED AND SA HEADER
568	1pfd		62	164	8e-28	0.15	0.33		FERREDOXIN; CHAIN: NULL;	ELECTRON TRANSPORT [2FE-2S] FERREDOXIN, SOLUTION STRUCTURE, PARAMAGNETISM, 2 NUCLEAR RELAXATION, ELECTRON TRANSPORT
568	1roe		62	166	4.8e-31	0.30	0.05		FERREDOXIN; CHAIN: NULL	ELECTRON TRANSPORT ELECTRON TRANSPORT, IRON-SULFUR
568	4fxc		62	166	8e-29	0.02	0.52		FERREDOXIN; 4FXC 4 CHAIN: NULL 4FXC 5	ELECTRON TRANSPORT
569	1avl	A	1	202	0.0009			53.78	APOLIPOPROTEIN A-I; CHAIN: A, B, C, D;	LIPID TRANSPORT APO A-I; LIPOPROTEIN, LIPID TRANSPORT, CHOLESTEROL METABOLISM, 2

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										ATHEROSCLEROSIS, HDL, LCAT-ACTIVATION
570	1c0a	A	1	502	0	0.58	1.00		ASPARTYL TRNA SYNTHETASE; CHAIN: A; ASPARTYL TRNA; CHAIN: B;	LIGASE/RNA ASPARTATE-TRNA LIGASE, ASPRS; PROTEIN-RNA COMPLEX
570	1g51	A	1	503	0	0.34	1.00		ASPARTYL-TRNA SYNTHETASE; CHAIN: A, B;	LIGASE AMINOACYL TRNA SYNTHETASE
571	1d0s	A	2	96	5.4e-09	0.28	-0.20		NICOTINATE MONONUCLEOTIDE:5,6- CHAIN: A;	TRANSFERASE DINUCLEOTIDE-BINDING MOTIF, PHOSPHORIBOSYL TRANSFERASE
571	1dx8	A	371	418	0.009	-0.64	0.09		RUBREDOXIN; CHAIN: A;	ELECTRON TRANSPORT NMR, RUBREDOXIN, GUILLARDIA THETA, ZINC-SUBSTITUTION
574	1qm4	A	16	257	0	0.71	1.00		METHIONINE ADENOSYLTRANSFERASE, ALPHA FORM; CHAIN: A, B;	TRANSFERASE ADOMET SYNTHETASE, MAT-I, ADENOSYLTRANSFERASE, METHIONINE BINDING
575	1chc		61	113	8e-12	0.47	0.41		VIRUS EQUINE HERPES VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	
575	1fbv	A	58	122	0.00036	-0.09	0.22		SIGNAL TRANSDUCTION PROTEIN CBL; CHAIN: A; ZAP-70 PEPTIDE; CHAIN: B; UBIQUITIN-CONJUGATING ENZYME E12-18 KDA UBCH7; CHAIN: C;	LIGASE CBL, UBCH7, ZAP-70, E2, UBIQUITIN, E3, PHOSPHORYLATION, 2 TYROSINE KINASE, UBIQUITINATION, PROTEIN DEGRADATION,
575	1rmd		63	121	1.6e-08	0.24	0.18		RAG1; CHAIN: NULL;	DNA-BINDING PROTEIN V(D)J RECOMBINATION ACTIVATING PROTEIN 1; RAG1, V(D)J RECOMBINATION, ANTIBODY, MAD, RING FINGER, 2 ZINC BINUCLEAR CLUSTER, ZINC FINGER, DNA-BINDING PROTEIN
581	1erj	A	91	371	4.8e-54	0.15	0.13		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
581	1fqv	A	81	126	1.8e-05	-0.12	0.35		SKP2; CHAIN: A, C, E, G, I, K, M,	LIGASE CYCLIN A/CDK2-ASSOCIATED

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
581	1fs1	A	81	120	0.00011	-0.37	0.65		CYCLIN A/CDK2-ASSOCIATED P19; CHAIN: A, C; CYCLIN A/CDK2-ASSOCIATED P45; CHAIN: B, D;	LIGASE SKP2 F-BOX; SKP1; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
581	1got	B	117	415	1.6e-54	0.29	0.93		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
581	1got	B	123	455	6.4e-71			96.38	GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
581	1got	B	162	455	6.4e-71	0.57	0.81		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
583	1b7f	A	172	302	9.6e-16	-0.00	-0.01		SXL-LETHAL PROTEIN; CHAIN: A, B; RNA (5'-R(P*GP*UP*UP*GP*UP*UP*UP*UP*UP*UP*UP*U)- CHAIN: P, Q;	RNA-BINDING PROTEIN/RNA TRA PRE-MRNA; SPLICING REGULATION, RNP DOMAIN, RNA COMPLEX
583	1b7f	A	2	135	3.2e-25	-0.10	0.09		SXL-LETHAL PROTEIN; CHAIN: A, B; RNA (5'-R(P*GP*UP*UP*GP*UP*UP*UP*U	RNA-BINDING PROTEIN/RNA TRA PRE-MRNA; SPLICING REGULATION, RNP DOMAIN, RNA COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									P*UP*UP*UP*U)- CHAIN: P, Q;	
583	1b7f	A	231	426	4.8e-34	-0.06	0.71		SXL-LETHAL PROTEIN; CHAIN: A, B; RNA (5'-R(P*GP*UP*UP*GP*UP*UP*UP*UP*UP*UP*U)- CHAIN: P, Q;	RNA-BINDING PROTEIN/RNA TRA PRE-MRNA; SPLICING REGULATION, RNP DOMAIN, RNA COMPLEX
583	1cvj	A	179	308	1.6e-21	0.21	0.09		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
583	1cvj	A	2	141	6.4e-28	-0.20	0.21		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
583	1cvj	A	232	432	4.8e-34	0.09	0.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
583	1cvj	A	66	204	8e-26	0.05	-0.11		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
583	1cvj	B	179	288	9.6e-16	-0.04	0.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
583	1cvj	B	2	121	8e-25	-0.29	0.47		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									R(*AP*AP*AP*AP*AP*AP*AP*AP *AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	REGULATION/RNA
583	1cvj	B	232	412	4.8e-29	0.10	0.46		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP *AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
583	1cvj	B	351	433	4.8e-18	0.25	-0.13		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP *AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
583	1cvj	B	66	202	3.2e-24	0.02	-0.12		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP *AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
583	1cvj	F	232	402	6.4e-20	-0.00	0.03		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP *AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
583	1cvj	F	351	433	4.8e-18	0.07	-0.06		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP *AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
583	1cvj	F	66	148	1.1e-22	0.21	0.31		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP *AP*AP*A)-3'); CHAIN: M, N, O, P,	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									Q, R, S, T;	
583	1cvj	H	232	405	1.4e-20	-0.05	0.29		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
583	1d8z	A	227	310	9.6e-16	-0.16	0.49		HU ANTIGEN C; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
583	1d8z	A	345	432	8e-19	-0.03	0.01		HU ANTIGEN C; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
583	1fht		224	328	6.4e-17	0.37	0.03		U1 SMALL NUCLEAR RIBONUCLEOPROTEIN A; CHAIN: NULL;	RIBONUCLEOPROTEIN U1A117; RIBONUCLEOPROTEIN, RNP DOMAIN, SPLICEOSOME
583	1ha1		172	302	4.8e-16	-0.25	0.12		HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
583	1ha1		2	135	8e-30	-0.10	0.01		HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
583	1ha1		65	210	4.8e-18	0.07	-0.13		HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
583	1hd1	A	231	302	3.2e-17	-0.04	0.15		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN D0; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
583	1hd1	A	350	426	1.6e-18	0.19	-0.19		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN D0; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
583	1qm9	A	227	428	3.2e-37	-0.33	0.93		POLYPYRIMIDINE TRACT-BINDING PROTEIN; CHAIN: A;	RIBONUCLEOPROTEIN PTB, PTB-C198, HETEROGENEOUS NUCLEAR POLYPYRIMIDINE TRACT BINDING

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										PROTEIN, RNP, RNA, SPICING, 2 TRANSLATION
583	1qm9	A	61	304	9.6e-18	-0.28	0.23		POLYPYRIMIDINE TRACT-BINDING PROTEIN; CHAIN: A;	RIBONUCLEOPROTEIN PTB, PTB-C198, HETEROGENEOUS NUCLEAR POLYPYRIMIDINE TRACT BINDING PROTEIN, RNP, RNA, SPICING, 2 TRANSLATION
583	2sxl		231	308	9.6e-16	0.26	1.00		SEX-LETHAL PROTEIN; CHAIN: NULL;	RNA-BINDING DOMAIN RNA-BINDING DOMAIN, ALTERNATIVE SPLICING
583	3sxl	A	2	125	8e-24	-0.57	0.07		SEX-LETHAL; CHAIN: A, B, C;	RNA BINDING DOMAIN RNA BINDING DOMAIN, RBD, RNA RECOGNITION MOTIF, RRM, 2 SPLICING INHIBITOR, TRANSLATIONAL INHIBITOR, SEX 3 DETERMINATION, X CHROMOSOME DOSAGE COMPENSATION
583	3sxl	A	231	419	3.2e-32	-0.12	0.58		SEX-LETHAL; CHAIN: A, B, C;	RNA BINDING DOMAIN RNA BINDING DOMAIN, RBD, RNA RECOGNITION MOTIF, RRM, 2 SPLICING INHIBITOR, TRANSLATIONAL INHIBITOR, SEX 3 DETERMINATION, X CHROMOSOME DOSAGE COMPENSATION
584	1sfc	A	25	93	3.2e-26	-0.41	0.90		SYNAPTOBREVIN 2; CHAIN: A, E, I; SYNTAXIN 1A; CHAIN: B, F, J; SNAP-25B; CHAIN: C, G, K; SNAP-25B; CHAIN: D, H, L;	TRANSPORT PROTEIN VAMP 2; MEMBRANE FUSION PROTEIN COMPLEX, TRANSPORT PROTEIN
584	1sfc	A	25	93	3.2e-26			106.16	SYNAPTOBREVIN 2; CHAIN: A, E, I; SYNTAXIN 1A; CHAIN: B, F, J; SNAP-25B; CHAIN: C, G, K; SNAP-25B; CHAIN: D, H, L;	TRANSPORT PROTEIN VAMP 2; MEMBRANE FUSION PROTEIN COMPLEX, TRANSPORT PROTEIN
586	1a9n	B	211	304	5.4e-05	-0.04	0.35		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
586	1cx0	A	211	276	1.8e-05	-0.18	0.33		U1A PROTEIN; CHAIN: A; HDV RIBOZYME SELF-CLEAVED; CHAIN: B;	RNA BINDING PROTEIN/RNA NESTED DOUBLE PSEUDOKNOT RNA STRUCTURE
586	1fht		211	276	7.2e-05	-0.17	0.22		U1 SMALL NUCLEAR	RIBONUCLEOPROTEIN U1A117;

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									RIBONUCLEOPROTEIN A; CHAIN: NULL;	RIBONUCLEOPROTEIN, RNP DOMAIN, SPLICEOSOME
586	1ha1		215	283	0.0018	-0.08	0.15		HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
586	1nrc	A	211	276	5.4e-05	-0.48	0.70		RIBONUCLEOPROTEIN PROTEIN FROM U1 SMALL NUCLEAR RIBONUCLEOPROTEIN (SNRNP U1) 1NRC 3 (N-TERMINAL FRAGMENT, RESIDUES 1 - 95) MUTANT WITH GLN 85 1NRC 4 REPLACED BY CYS (Q85C) 1NRC 5	
586	1urn	A	210	276	1.8e-05	0.25	0.89		U1A SPLICEOSOMAL PROTEIN; 1URN 5 CHAIN: A, B, C; 1URN 6 RNA 21MER HAIRPIN (5'-(AP*AP*UP*CP*CP*AP*UP*UP* 1URN 11 CHAIN: P, Q, R 1URN 13	COMPLEX (RIBONUCLEOPROTEIN/RNA)
586	2u1a		215	298	0.00072	-0.23	0.19		U1 SMALL NUCLEAR RIBONUCLEOPROTEIN A; CHAIN: NULL;	NUCLEAR PROTEIN U1 SNRNP A PROTEIN; RNA BINDING DOMAIN, NUCLEAR PROTEIN
586	3sxl	A	215	276	0.0054	-0.15	0.03		SEX-LETHAL; CHAIN: A, B, C;	RNA BINDING DOMAIN RNA BINDING DOMAIN, RBD, RNA RECOGNITION MOTIF, RRM, 2 SPLICING INHIBITOR, TRANSLATIONAL INHIBITOR, SEX 3 DETERMINATION, X CHROMOSOME DOSAGE COMPENSATION
589	1aj4		70	227	6.4e-44			116.57	TROPONIN C; CHAIN: NULL;	MUSCLE PROTEIN CTNC; CARDIAC, MUSCLE PROTEIN, REGULATORY, CALCIUM BINDING
589	1aj4		99	226	6.4e-44	0.55	1.00		TROPONIN C; CHAIN: NULL;	MUSCLE PROTEIN CTNC; CARDIAC, MUSCLE PROTEIN, REGULATORY, CALCIUM BINDING
589	1aj5	A	17	192	3.6e-22			56.61	CALPAIN; CHAIN: A, B;	CALCIUM-BINDING PROTEIN CALCIUM-BINDING PROTEIN, CALCIUM-

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										DEPENDENT PROTEASE, APO 2 FORM, SMALL SUBUNIT
589	1aj5	A	82	186	3.6e-22	0.24	0.98		CALPAIN; CHAIN: A, B;	CALCIUM-BINDING PROTEIN CALCIUM-BINDING PROTEIN, CALCIUM-DEPENDENT PROTEASE, APO 2 FORM, SMALL SUBUNIT
589	1ak8		65	150	1.3e-23	-0.21	0.37		CALMODULIN; CHAIN: NULL;	CALCIUM-BINDING PROTEIN CALMODULIN CERIUM TR1C-DOMAIN, RESIDUES 1 - 75; CERIUM-LOADED, CALCIUM-BINDING PROTEIN
589	1alv	A	8	192	1.4e-20			57.63	CALPAIN; CHAIN: A, B;	CALCIUM BINDING S-CAMLD; CALCIUM BINDING, CALMODULIN LIKE, DOMAIN OF CYSTEIN 2 PROTEASE
589	1alv	A	82	186	1.4e-20	0.35	1.00		CALPAIN; CHAIN: A, B;	CALCIUM BINDING S-CAMLD; CALCIUM BINDING, CALMODULIN LIKE, DOMAIN OF CYSTEIN 2 PROTEASE
589	1ap4		79	151	5.4e-20	0.97	1.00		CARDIAC N-TROPONIN C; CHAIN: NULL;	CALCIUM-BINDING CNTNC; CALCIUM-BINDING, REGULATION, TROPONIN C, CARDIAC MUSCLE 2 CONTRACTION
589	1aui	B	70	227	1.4e-24			84.92	SERINE/THREONINE PHOSPHATASE 2B; CHAIN: A, B;	HYDROLASE CALCINEURIN; HYDROLASE, PHOSPHATASE, IMMUNOSUPPRESSION
589	1bjf	A	65	227	1.6e-18			59.41	NEUROCALCIN DELTA; CHAIN: A, B;	CALCIUM-BINDING CALCIUM-BINDING, MYRISTOYLATION, NEURONAL SPECIFIC GUANYLATE 2 CYCLASE ACTIVATOR
589	1cdm	A	79	226	6.4e-54			128.94	CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING DOMAIN OF 1CDM 3 CALMODULIN-DEPENDENT PROTEIN KINASE II 1CDM 4	
589	1cdm	A	94	226	6.4e-54	0.59	1.00		CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING DOMAIN OF 1CDM 3 CALMODULIN-DEPENDENT PROTEIN KINASE II 1CDM 4	
589	1cll		79	227	6.4e-59			142.28	CALCIUM-BINDING PROTEIN	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CALMODULIN (VERTEBRATE) 1CLL 3	
589	1cll		94	226	6.4e-59	0.43	1.00		CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) 1CLL 3	
589	1cmf		155	227	7.2e-21			69.28	CALMODULIN (VERTEBRATE); 1CMF 6 CHAIN: NULL; 1CMF 7	CALCIUM-BINDING PROTEIN CALMODULIN APO TR2C-DOMAIN; 1CMF 9
589	1df0	A	84	182	5.4e-22	0.12	0.99		M-CALPAIN; CHAIN: A; CALPAIN; CHAIN: B;	HYDROLASE CALCIUM-ACTIVATED NEUTRAL PROTEINASE, CALPAIN 2; CALCIUM-ACTIVATED NEUTRAL PROTEINASE; CYSTEINE PROTEASE, CALMODULIN, PAPAIN, CATALYTIC TRIAD, 2 ZYMOGEN ACTIVATION, CALCIUM, C2 DOMAIN, PROTEASE, ZYMOGEN, 3 CALPAIN
589	1dkv	A	84	180	3.6e-21	0.18	0.96		M-CALPAIN; CHAIN: A; CALPAIN; CHAIN: B;	HYDROLASE CALCIUM-ACTIVATED NEUTRAL PROTEINASE; CALCIUM-ACTIVATED NEUTRAL PROTEINASE; M-CALPAIN, CALCIUM, PAPAIN-LIKE
589	1dkv	B	82	186	3.6e-22	0.47	0.99		M-CALPAIN; CHAIN: A; CALPAIN; CHAIN: B;	HYDROLASE CALCIUM-ACTIVATED NEUTRAL PROTEINASE; CALCIUM-ACTIVATED NEUTRAL PROTEINASE; M-CALPAIN, CALCIUM, PAPAIN-LIKE
589	1exr	A	67	225	3.2e-57	0.37	1.00		CALMODULIN; CHAIN: A;	METAL TRANSPORT CALMODULIN, HIGH RESOLUTION, DISORDER
589	1tcf		70	227	1.4e-45			127.75	TROPONIN C; CHAIN: NULL;	CALCIUM-REGULATED MUSCLE CONTRACTION MUSCLE CONTRACTION, CALCIUM-BINDING, TROPONIN, E-F HAND, 2 OPEN CONFORMATION REGULATORY DOMAIN, CALCIUM-REGULATED 3 MUSCLE CONTRACTION
589	1tcf		94	225	1.4e-45	0.16	1.00		TROPONIN C; CHAIN: NULL;	CALCIUM-REGULATED MUSCLE CONTRACTION MUSCLE CONTRACTION, CALCIUM-BINDING, TROPONIN, E-F HAND, 2 OPEN CONFORMATION REGULATORY DOMAIN, CALCIUM-REGULATED 3 MUSCLE CONTRACTION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
589	1tnx		70	225	1.1e-41			116.06	TROPONIN C; 1TNX 4 CHAIN: NULL; 1TNX 5	CALCIUM-BINDING PROTEIN EF-HAND 1TNX 14
589	1top		67	227	1.6e-46			125.84	CONTRACTILE SYSTEM PROTEIN TROPONIN C 1TOP 3	
589	1top		94	225	1.6e-46	0.64	1.00		CONTRACTILE SYSTEM PROTEIN TROPONIN C 1TOP 3	
589	1trc	A	160	227	3.6e-21			67.46	CALCIUM BINDING PROTEIN CALMODULIN (/TR=2=C$ FRAGMENT COMPRISING RESIDUES 78 - 148 1TRC 3 OF THE INTACT MOLECULE) 1TRC 4	
589	1vrk	A	66	227	1.6e-57	0.42	1.00		CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALLING, 2 COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
589	1vrk	A	76	227	1.6e-57			139.36	CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALLING, 2 COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
589	1wdc	C	79	227	3.6e-40			90.57	SCALLOP MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MYOSIN, CALCIUM BINDING PROTEIN, MUSCLE PROTEIN
590	1a88	A	529	771	1.6e-07	-0.16	0.28		CHLOROPEROXIDASE L; CHAIN: A, B, C;	HALOPEROXIDASE BROMOPEROXIDASE L, HALOPEROXIDASE L; HALOPEROXIDASE, OXIDOREDUCTASE
590	1a8s		528	771	4.8e-09	-0.31	0.52		CHLOROPEROXIDASE F; CHAIN: NULL;	HALOPEROXIDASE HALOPEROXIDASE F; HALOPEROXIDASE, OXIDOREDUCTASE, PROPIONATE COMPLEX
590	1auo	A	556	775	8e-12	-0.04	0.43		CARBOXYLESTERASE; CHAIN: A, B;	HYDROLASE HYDROLASE
590	1c4x	A	551	776	9.6e-08	-0.06	0.11		2-HYDROXY-6-OXO-6-PHENYLHEXA-2,4-DIENOATE CHAIN: A;	HYDROLASE BPHD; HYDROLASE, PCB DEGRADATION
590	1cle	A	483	780	1.6e-56	-0.36	0.07		CHOLESTEROL ESTERASE; 1CLE 4 CHAIN: A, B; 1CLE 5	LIPASE ESTERASE, SUBSTRATE/PRODUCT-BOUND 1CLE 9
590	1crz	A	177	453	0.0036	0.13	0.07		TOLB PROTEIN; CHAIN: A;	TOXIN BINDING PROTEIN TWO DOMAINS: BETA PROPELLER AND ALPHA/BETA FOLD

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
590	1dx4	A	461	764	1.6e-65	-0.37	0.28		ACETYLCHOLINESTERASE; CHAIN: A;	HYDROLASE (SERINE ESTERASE) HYDROLASE (SERINE ESTERASE), HYDROLASE, SERINE ESTERASE, 2 SYNAPSE, MEMBRANE, NERVE, MUSCLE, SIGNAL, NEUROTRANSMITTER 3 DEGRADATION, GLYCOPROTEIN, GPI-ANCHOR, ALTERNATIVE SPLICING
590	1ea5	A	463	760	1.3e-68	-0.22	0.17		ACETYLCHOLINESTERASE; CHAIN: A;	CHOLINESTERASE SERINE HYDROLASE, NEUROTRANSMITTER CLEAVAGE, CATALYTIC 2 TRIAD, ALPHA/BETA HYDROLASE
590	1ehy	A	713	770	0.00036	-0.38	0.16		SOLUBLE EPOXIDE HYDROLASE; CHAIN: A, B, C, D;	HYDROLASE HYDROLASE, ALPHA/BETA HYDROLASE FOLD, EPOXIDE DEGRADATION, 2 EPICHLOROHYDRIN
590	1evq	A	519	774	1.6e-36	-0.11	0.87		SERINE HYDROLASE; CHAIN: A;	HYDROLASE ALPHA/BETA HYDROLASE FOLD
590	1fj2	A	712	766	3.6e-06	0.16	0.35		ACYL PROTEIN THIOESTERASE 1; CHAIN: A, B;	HYDROLASE ALPHA/BETA HYDROLASE, SERINE HYDROLASE, SAD, ANOMALOUS 2 DIFFRACTION
590	1jkm	A	529	768	1.6e-22	0.17	0.68		BREFELDIN A ESTERASE; CHAIN: A, B;	SERINE HYDROLASE SERINE HYDROLASE, DEGRADATION OF BREFELDIN A, ALPHA/BETA 2 HYDROLASE FAMILY
590	1lpp		483	780	4.8e-55	-0.29	0.03		HYDROLASE LIPASE (E.C.3.1.1.3) (TRIACYLGLYCEROL LIPASE) COMPLEXED WITH 1LPP 3 HEXADECANESULFONATE 1LPP 4 1LPP 71	
590	1maa	A	463	760	9.6e-70	-0.22	0.21		ACETYLCHOLINESTERASE; CHAIN: A, B, C, D;	HYDROLASE MACHE; HYDROLASE, SERINE ESTERASE, ACETYLCHOLINESTERASE, TETRAMER, 2 HYDROLASE FOLD, GLYCOSYLATED PROTEIN
590	1qe3	A	462	779	1.6e-67	-0.26	0.33		PARA-NITROBENZYL ESTERASE; CHAIN: A;	HYDROLASE PNB ESTERASE; ALPHA-BETA HYDROLASE DIRECTED EVOLUTION
590	1qfm	A	174	777	9.6e-80	-0.00	0.90		PROLYL OLIGOPEPTIDASE; CHAIN: A;	HYDROLASE PROLYL ENDOPEPTIDASE, POST-PROLINE CLEAVING PROLYL

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										OLIGOPEPTIDASE, AMNESIA, ALPHA/BETA-HYDROLASE, BETA-2 PROPELLER
590	1qfm	A	72	782	5.4e-84			125.95	PROLYL OLIGOPEPTIDASE; CHAIN: A;	HYDROLASE PROLYL ENDOPEPTIDASE, POST-PROLINE CLEAVING PROLYL OLIGOPEPTIDASE, AMNESIA, ALPHA/BETA-HYDROLASE, BETA-2 PROPELLER
590	1qfm	A	80	778	5.4e-84	0.09	0.94		PROLYL OLIGOPEPTIDASE; CHAIN: A;	HYDROLASE PROLYL ENDOPEPTIDASE, POST-PROLINE CLEAVING PROLYL OLIGOPEPTIDASE, AMNESIA, ALPHA/BETA-HYDROLASE, BETA-2 PROPELLER
590	1qtr	A	520	699	6.4e-07	-0.02	0.05		PROLYL AMINOPEPTIDASE; CHAIN: A;	HYDROLASE ALPHA BETA HYDROLASE FOLD, PROLINE, PROLYL AMINOPEPTIDASE, 2 SERRATIA, IMINOPEPTIDASE
590	1thg		473	774	1.6e-57	-0.22	0.31		HYDROLASE(CARBOXYLIC ESTERASE) LIPASE (E.C.3.1.1.3) TRIACYLGLYCEROL HYDROLASE 1THG 3	
592	1ddv	A	17	121	1.8e-19	0.62	0.94		GLGF-DOMAIN PROTEIN HOMER; CHAIN: A; METABOTROPIC GLUTAMATE RECEPTOR MGLUR5; CHAIN: B;	SIGNALING PROTEIN PROTEIN-LIGAND COMPLEX, POLYPROLINE RECOGNITION, BETA TURN
592	1ddv	A	47	113	0.0018	0.30	0.06		GLGF-DOMAIN PROTEIN HOMER; CHAIN: A; METABOTROPIC GLUTAMATE RECEPTOR MGLUR5; CHAIN: B;	SIGNALING PROTEIN PROTEIN-LIGAND COMPLEX, POLYPROLINE RECOGNITION, BETA TURN
592	1ddw	A	11	113	0.00032	0.15	0.07		GLGF-DOMAIN PROTEIN HOMER; CHAIN: A;	SIGNALING PROTEIN PLECKSTRIN HOMOLOGY DOMAIN FOLD
592	1ddw	A	17	121	1.8e-18	0.72	0.99		GLGF-DOMAIN PROTEIN HOMER; CHAIN: A;	SIGNALING PROTEIN PLECKSTRIN HOMOLOGY DOMAIN FOLD
592	1egx	A	10	125	3.2e-40	0.31	0.69		VASODILATOR-STIMULATED PHOSPHOPROTEIN; CHAIN: A;	SIGNALING PROTEIN VASP; EVH1, VASP-ENA, NMR, POLY-PROLINE-BINDING DOMAIN
592	1evh	A	10	123	1.4e-43	0.36	0.66		MENA EVH1 DOMAIN; CHAIN: A;	CONTRACTILE PROTEIN WH1 DOMAIN;

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									PEPTIDE ACTA; CHAIN: B;	MOLECULAR RECOGNITION, ACTIN DYNAMICS, CONTRACTILE PROTEIN
592	1evh	A	10	124	1.4e-43			85.76	MENA EVH1 DOMAIN; CHAIN: A; PEPTIDE ACTA; CHAIN: B;	CONTRACTILE PROTEIN WH1 DOMAIN; MOLECULAR RECOGNITION, ACTIN DYNAMICS, CONTRACTILE PROTEIN
592	1qc6	A	10	121	6.4e-40	0.37	0.95		EVH1 DOMAIN FROM ENA/VASP-LIKE PROTEIN; CHAIN: A, B; PHE-GLU-PHE-PRO-PRO-PRO-PRO-THR-ASP-GLU-GLU; CHAIN: C, D;	CELL MOTILITY AN INCOMPLETE SEVEN STRANDED ANTI-PARALLEL BETA BARREL 2 CLOSED BY AN ALPHA HELIX, EVH1 DOMAIN, ACTIN-BASED CELL 3 MOTILITY, INTERACTION MODULE
592	1qc6	A	10	122	6.4e-40			60.92	EVH1 DOMAIN FROM ENA/VASP-LIKE PROTEIN; CHAIN: A, B; PHE-GLU-PHE-PRO-PRO-PRO-PRO-THR-ASP-GLU-GLU; CHAIN: C, D;	CELL MOTILITY AN INCOMPLETE SEVEN STRANDED ANTI-PARALLEL BETA BARREL 2 CLOSED BY AN ALPHA HELIX, EVH1 DOMAIN, ACTIN-BASED CELL 3 MOTILITY, INTERACTION MODULE
593	1a6q		182	536	9.6e-45	0.10	-0.07		PHOSPHATASE 2C; CHAIN: NULL;	HYDROLASE CATALYTIC MECHANISM, METALLOENZYME, PROTEIN PHOSPHATASE 2C, 2 SIGNAL TRANSDUCTUIN, X-RAY CRYSTALLOGRAPHY, HYDROLASE
595	1crz	A	15	256	3.2e-05	0.09	0.42		TOLB PROTEIN; CHAIN: A;	TOXIN BINDING PROTEIN TWO DOMAINS: BETA PROPELLER AND ALPHA/BETA FOLD
595	1erj	A	1	231	1.6e-48	0.06	0.87		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
595	1erj	A	3	280	3.6e-17	0.19	1.00		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
595	1erj	A	51	359	1.3e-69	0.15	0.83		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
595	1got	B	1	273	1.6e-53	0.47	0.89		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
595	1got	B	2	356	1.3e-80			80.19	GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
595	1got	B	41	356	1.3e-80	0.27	0.87		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
596	1am4	D	5	173	3.6e-56			80.47	P50-RHOGAP; CHAIN: A, B, C; CDC42HS; CHAIN: D, E, F;	COMPLEX (GTPASE-ACTIVATING/GTP-BINDING) COMPLEX (GTPASE-ACTIVATING/GTP-BINDING), GTPASE ACTIVATION
596	1byu	A	3	200	1.8e-59			105.61	GTP-BINDING PROTEIN RAN; CHAIN: A, B;	TRANSPORT PROTEIN TC4; GTPASE, NUCLEAR TRANSPORT, TRANSPORT PROTEIN
596	1byu	A	7	178	1.8e-59	0.44	1.00		GTP-BINDING PROTEIN RAN; CHAIN: A, B;	TRANSPORT PROTEIN TC4; GTPASE, NUCLEAR TRANSPORT, TRANSPORT PROTEIN
596	1byu	B	1	200	1.1e-59			101.37	GTP-BINDING PROTEIN RAN; CHAIN: A, B;	TRANSPORT PROTEIN TC4; GTPASE, NUCLEAR TRANSPORT, TRANSPORT PROTEIN
596	1byu	B	2	178	1.1e-59	0.52	1.00		GTP-BINDING PROTEIN RAN; CHAIN: A, B;	TRANSPORT PROTEIN TC4; GTPASE, NUCLEAR TRANSPORT, TRANSPORT PROTEIN
596	1cly	A	6	172	3.2e-65			95.03	RAS-RELATED PROTEIN RAP-1A; CHAIN: A; PROTO-ONKOGENE SERINE/THREONINE PROTEIN KINASE CHAIN: B;	SIGNALING PROTEIN GTP-BINDING PROTEINS, PROTEIN-PROTEIN COMPLEX, EFFECTORS
596	1cly	A	6	174	3.2e-65	0.62	1.00		RAS-RELATED PROTEIN RAP-1A; CHAIN: A; PROTO-ONKOGENE SERINE/THREONINE PROTEIN	SIGNALING PROTEIN GTP-BINDING PROTEINS, PROTEIN-PROTEIN COMPLEX, EFFECTORS

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									KINASE CHAIN: B;	
596	1ctq	A	6	174	1.1e-65			101.74	TRANSFORMING PROTEIN P21/H-RAS-1; CHAIN: A;	SIGNALING PROTEIN G PROTEIN, GTP HYDROLYSIS, KINETIC CRYSTALLOGRAPHY, 2 SIGNALING PROTEIN
596	1ctq	A	7	175	1.1e-65	0.90	1.00		TRANSFORMING PROTEIN P21/H-RAS-1; CHAIN: A;	SIGNALING PROTEIN G PROTEIN, GTP HYDROLYSIS, KINETIC CRYSTALLOGRAPHY, 2 SIGNALING PROTEIN
596	1cxz	A	2	175	8e-58			91.18	HIS-TAGGED TRANSFORMING PROTEIN RHOA(0-181); CHAIN: A; PKN; CHAIN: B;	SIGNALING PROTEIN PROTEIN-PROTEIN COMPLEX, ANTIPARALLEL COILED-COIL
596	1d5c	A	9	172	1.8e-60	0.86	1.00		RAB6 GTPASE; CHAIN: A;	ENDOCYTOSIS/EXOCYTOSIS G-PROTEIN, GTPASE, RAB6, VESICULAR TRAFFICKING
596	1e0s	A	7	141	6.4e-15	0.22	0.13		ADP-RIBOSYLATION FACTOR 6; CHAIN: A;	G PROTEIN G PROTEIN, RAS, ARF, ARF6, MEMBRANE TRAFFIC
596	1ibr	A	6	181	1.3e-60	0.69	1.00		RAN; CHAIN: A, C; IMPORTIN BETA SUBUNIT; CHAIN: B, D;	SMALL GTPASE KARYOPHERIN BETA, P95 SMALL GTPASE, NUCLEAR TRANSPORT RECEPTOR
596	1ibr	A	6	181	1.3e-60			116.07	RAN; CHAIN: A, C; IMPORTIN BETA SUBUNIT; CHAIN: B, D;	SMALL GTPASE KARYOPHERIN BETA, P95 SMALL GTPASE, NUCLEAR TRANSPORT RECEPTOR
596	1kao		6	175	9.6e-59			101.06	RAP2A; CHAIN: NULL;	GTP-BINDING PROTEIN GTP-BINDING PROTEIN, SMALL G PROTEIN, RAP2, GDP, RAS
596	1mh1		3	177	3.2e-59			93.80	RAC1; CHAIN: NULL;	GTP-BINDING GTP-BINDING, GTPASE, SMALL G-PROTEIN, RHO FAMILY, RAS SUPER 2 FAMILY
596	1mh1		8	177	3.2e-59	0.57	1.00		RAC1; CHAIN: NULL;	GTP-BINDING GTP-BINDING, GTPASE, SMALL G-PROTEIN, RHO FAMILY, RAS SUPER 2 FAMILY
596	1plj		8	174	9e-49			53.73	ONCOGENE PROTEIN C-H-RAS P21 PROTEIN MUTANT WITH GLY 12 REPLACED BY PRO 1PLJ 3 (G12P) COMPLEXED WITH P3-1-(2-NITROPHENYL)ETHYL- 1PLJ 4 GUANOSINE-5'-(B,G-IMIDO)-	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									TRIPHOSPHATE 1PLJ 5	
596	1rrp	C	5	192	7.2e-60			117.23	RAN; CHAIN: A, C; NUCLEAR PORE COMPLEX PROTEIN NUP358; CHAIN: B, D;	COMPLEX (SMALL GTPASE/NUCLEAR PROTEIN) COMPLEX (SMALL GTPASE/NUCLEAR PROTEIN), SMALL GTPASE, 2 NUCLEAR TRANSPORT
596	1rrp	C	6	178	7.2e-60	0.61	1.00		RAN; CHAIN: A, C; NUCLEAR PORE COMPLEX PROTEIN NUP358; CHAIN: B, D;	COMPLEX (SMALL GTPASE/NUCLEAR PROTEIN) COMPLEX (SMALL GTPASE/NUCLEAR PROTEIN), SMALL GTPASE, 2 NUCLEAR TRANSPORT
596	1tx4	B	5	173	9e-57			86.19	P50-RHOGAP; CHAIN: A; TRANSFORMING PROTEIN RHOA; CHAIN: B;	COMPLEX(GTPASE ACTIVATN/PROTO-ONCOGENE) GTPASE-ACTIVATING PROTEIN RHOGAP; COMPLEX (GTPASE ACTIVATION/PROTO-ONCOGENE), GTPASE, 2 TRANSITION STATE, GAP
596	1zbd	A	3	180	8e-65			110.73	RAB-3A; CHAIN: A; RABPHILIN-3A; CHAIN: B;	COMPLEX (GTP-BINDING/EFFECTOR) RAS-RELATED PROTEIN RAB3A; COMPLEX (GTP-BINDING/EFFECTOR), G PROTEIN, EFFECTOR, RABCDR, 2 SYNAPTIC EXOCYTOSIS, RAB PROTEIN, RAB3A, RABPHILIN
596	1zbd	A	4	178	8e-65	0.57	1.00		RAB-3A; CHAIN: A; RABPHILIN-3A; CHAIN: B;	COMPLEX (GTP-BINDING/EFFECTOR) RAS-RELATED PROTEIN RAB3A; COMPLEX (GTP-BINDING/EFFECTOR), G PROTEIN, EFFECTOR, RABCDR, 2 SYNAPTIC EXOCYTOSIS, RAB PROTEIN, RAB3A, RABPHILIN
596	2ngr	A	6	184	1.6e-57			83.18	GTP BINDING PROTEIN (G25K); CHAIN: A; GTPASE ACTIVATING PROTEIN (RHG); CHAIN: B;	HYDROLASE CDC42/CDC42GAP; CDC42/CDC42GAP; TRANSITION STATE, G-PROTEIN, GAP, CDC42, ALF3., HYDROLASE
596	3rab	A	3	175	1.1e-65	0.72	1.00		RAB3A; CHAIN: A;	HYDROLASE G PROTEIN, VESICULAR TRAFFICKING, GTP HYDROLYSIS, RAB 2 PROTEIN, NEUROTRANSMITTER RELEASE, HYDROLASE
596	3rab	A	4	174	1.1e-65			116.82	RAB3A; CHAIN: A;	HYDROLASE G PROTEIN, VESICULAR TRAFFICKING, GTP HYDROLYSIS, RAB 2 PROTEIN, NEUROTRANSMITTER RELEASE, HYDROLASE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
598	1mey	C	146	227	6.4e-47	0.49	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
598	1mey	C	146	227	9e-49	0.49	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
598	1mey	C	146	228	9e-49			109.88	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
598	1mey	C	174	255	4.8e-48	0.61	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
598	1mey	C	202	283	6.4e-49	0.01	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
598	1mey	C	230	311	6.4e-50	0.35	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
598	1mey	C	258	339	3.2e-50	0.42	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
598	1mey	C	2	60	4.8e-28	-0.37	0.01		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									PROTEIN; CHAIN: C, F, G;	PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
598	1mey	C	286	367	1.6e-50	0.38	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
598	1mey	C	342	423	1.4e-50	0.59	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
598	1mey	C	370	451	3.2e-50	0.37	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
598	1mey	C	398	479	1.6e-50	0.12	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
598	1mey	C	426	487	6.4e-38	0.12	0.93		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
598	1mey	C	63	143	6.4e-41	-0.30	0.37		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
598	1mey	C	90	171	8e-45	0.19	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
598	1tf6	A	175	320	1.6e-37	-0.05	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
598	1tf6	A	314	482	4.8e-38			115.02	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
598	1tf6	A	343	487	4.8e-37	-0.04	1.00		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
598	1tf6	A	64	208	4.8e-34	-0.12	0.78		TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
598	1ubd	C	144	255	1.8e-56	0.26	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
598	1ubd	C	146	256	1.8e-56			93.55	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
598	1ubd	C	182	283	3.2e-34	0.36	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
598	1ubd	C	200	311	1.8e-55	0.30	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
598	1ubd	C	210	311	6.4e-35	-0.28	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
598	1ubd	C	228	340	3.6e-55	-0.09	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
598	1ubd	C	238	339	1.1e-35	-0.06	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
598	1ubd	C	256	367	1.3e-56	0.17	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
598	1ubd	C	312	423	9e-55	0.26	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
598	1ubd	C	341	451	3.6e-53	0.05	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
598	1ubd	C	350	451	8e-35	0.11	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
598	1ubd	C	368	480	3.6e-51	0.25	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
598	1ubd	C	378	479	9.6e-35	0.04	0.95		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
598	1ubd	C	38	143	1.6e-28	-0.53	0.19		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: A, B;	ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
598	1ubd	C	71	171	1.4e-29	-0.35	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
598	1ubd	C	90	199	3.6e-49	0.01	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
598	2gli	A	118	257	1.8e-68	0.23	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
598	2gli	A	118	257	1.8e-68			98.70	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
598	2gli	A	174	369	3.6e-73	-0.32	0.70		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
598	2gli	A	266	394	1.6e-34	0.30	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
598	2gli	A	286	425	1.4e-70	0.15	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
598	2gli	A	342	481	5.4e-68	0.22	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
598	2gli	A	350	481	4.8e-34	0.13	0.95		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										COMPLEX (DNA-BINDING PROTEIN/DNA)
598	2gli	A	370	488	7.2e-40	-0.14	0.34		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
598	2gli	A	378	487	3.2e-29	-0.15	0.35		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
598	2gli	A	63	198	8e-30	0.23	0.54		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
598	2gli	A	71	201	3.6e-50	0.03	0.94		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
599	1a8l		36	260	4.8e-24			55.20	PROTEIN DISULFIDE OXIDOREDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE OXIDOREDUCTASE, PDI, THIOREDOXIN FOLD
599	1a8l		51	259	4.8e-24	0.19	0.47		PROTEIN DISULFIDE OXIDOREDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE OXIDOREDUCTASE, PDI, THIOREDOXIN FOLD
599	1a8y		31	266	5.4e-22	0.23	0.21		CALSEQUESTRIN; CHAIN: NULL	CALCIUM-BINDING PROTEIN CALSEQUESTRIN, CALCIUM-BINDING PROTEIN, SARCOPLASMIC 2 RETICULUM, RABBIT SKELETAL MUSCLE
599	1bjx		10	49	0.00013	-0.36	0.30		PROTEIN DISULFIDE ISOMERASE; CHAIN: NULL;	ELECTRON TRANSPORT ELECTRON TRANSPORT, REDOX-ACTIVE CENTER, ISOMERASE, 2 ENDOPLASMIC RETICULUM
599	1cqg	A	41	136	3.2e-22	0.65	0.13		THIOREDOXIN; CHAIN: A; REF-1 PEPTIDE; CHAIN: B;	COMPLEX (ELECTRON TRANSPORT/PEPTIDE) COMPLEX, ELECTRON TRANSPORT/PEPTIDE
599	1dby	A	155	264	9.6e-28			60.83	CHLOROPLAST THIOREDOXIN M CH2; CHAIN: A;	OXIDOREDUCTASE THIOREDOXIN M, THIOREDOXIN CH2, CHLOROPLASTIC THIOREDOXIN
599	1dby	A	162	260	9.6e-28	0.66	0.99		CHLOROPLAST THIOREDOXIN M CH2; CHAIN: A;	OXIDOREDUCTASE THIOREDOXIN M, THIOREDOXIN CH2, CHLOROPLASTIC THIOREDOXIN
599	1dby	A	2	39	9e-05	-0.02	0.47		CHLOROPLAST THIOREDOXIN M	OXIDOREDUCTASE THIOREDOXIN M,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CH2; CHAIN: A;	THIOREDOXIN CH2, CHLOROPLASTIC THIOREDOXIN
599	1dby	A	45	136	9.6e-26	0.81	0.64		CHLOROPLAST THIOREDOXIN M CH2; CHAIN: A;	OXIDOREDUCTASE THIOREDOXIN M, THIOREDOXIN CH2, CHLOROPLASTIC THIOREDOXIN
599	1erv		153	261	1.6e-22	0.20	0.48		THIOREDOXIN; CHAIN: NULL;	OXIDOREDUCTASE DIMER, THIOREDOXIN, X-RAY CRYSTALLOGRAPHY, OXIDOREDUCTASE
599	1erv		41	136	8e-25	0.26	0.58		THIOREDOXIN; CHAIN: NULL;	OXIDOREDUCTASE DIMER, THIOREDOXIN, X-RAY CRYSTALLOGRAPHY, OXIDOREDUCTASE
599	1f9m	A	148	257	9.6e-19	-0.01	0.47		THIOREDOXIN F; CHAIN: A, B;	ELECTRON TRANSPORT ELECTRON TRANSPORT
599	1f9m	A	39	132	6.4e-21	0.36	0.72		THIOREDOXIN F; CHAIN: A, B;	ELECTRON TRANSPORT ELECTRON TRANSPORT
599	1faa	A	148	257	9.6e-19	0.40	0.55		THIOREDOXIN F; CHAIN: A;	ELECTRON TRANSPORT ELECTRON TRANSPORT
599	1fb6	A	164	259	4.8e-29	0.62	0.95		THIOREDOXIN M; CHAIN: A, B;	ELECTRON TRANSPORT ELECTRON TRANSPORT
599	1fb6	A	2	38	0.00018	-0.50	0.10		THIOREDOXIN M; CHAIN: A, B;	ELECTRON TRANSPORT ELECTRON TRANSPORT
599	1fb6	A	42	146	6.4e-27	0.46	0.99		THIOREDOXIN M; CHAIN: A, B;	ELECTRON TRANSPORT ELECTRON TRANSPORT
599	1mek		152	267	3.2e-23			68.70	PROTEIN DISULFIDE ISOMERASE; CHAIN: NULL;	ELECTRON TRANSPORT ELECTRON TRANSPORT, REDOX-ACTIVE CENTER, ISOMERASE, 2 ENDOPLASMIC RETICULUM
599	1mek		165	267	3.2e-23	0.42	0.96		PROTEIN DISULFIDE ISOMERASE; CHAIN: NULL;	ELECTRON TRANSPORT ELECTRON TRANSPORT, REDOX-ACTIVE CENTER, ISOMERASE, 2 ENDOPLASMIC RETICULUM
599	1qgv	A	159	275	1.4e-21	0.39	0.31		SPLICEOSOMAL PROTEIN U5-15KD; CHAIN: A;	TRANSCRIPTION SPLICEOSOMAL PROTEIN, SNRNP, THIOREDOXIN, TRANSCRIPTION
599	1quw	A	157	259	3.2e-27	0.53	1.00		THIOREDOXIN; CHAIN: A;	ELECTRON TRANSPORT ALPHA/BETA OPEN-TWISTED PROTEIN, THIOL-DISULFIDE
599	1quw	A	2	38	0.00014	-0.08	0.01		THIOREDOXIN; CHAIN: A;	ELECTRON TRANSPORT ALPHA/BETA

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										OPEN-TWISTED PROTEIN, THIOL-DISULFIDE
599	1quw	A	45	139	3.2e-26	0.28	0.34		THIOREDOXIN; CHAIN: A;	ELECTRON TRANSPORT ALPHA/BETA OPEN-TWISTED PROTEIN, THIOL-DISULFIDE
599	1t7p	B	156	263	3.2e-29			63.78	DNA POLYMERASE; CHAIN: A; THIOREDOXIN; CHAIN: B; DNA; CHAIN: P, T;	T7 DNA POLYMERASE, DNA REPLICATION, NUCLEOTIDYL 2 TRANSFERASE, SEQUENCING, THIOREDOXIN, PROCESSIVITY FACTOR, 3 COMPLEX (HYDROLASE/ELECTRON TRANSPORT/DNA)
599	1t7p	B	158	259	3.2e-29	0.61	0.93		DNA POLYMERASE; CHAIN: A; THIOREDOXIN; CHAIN: B; DNA; CHAIN: P, T;	T7 DNA POLYMERASE, DNA REPLICATION, NUCLEOTIDYL 2 TRANSFERASE, SEQUENCING, THIOREDOXIN, PROCESSIVITY FACTOR, 3 COMPLEX (HYDROLASE/ELECTRON TRANSPORT/DNA)
599	1t7p	B	2	35	0.00018	-0.58	0.10		DNA POLYMERASE; CHAIN: A; THIOREDOXIN; CHAIN: B; DNA; CHAIN: P, T;	T7 DNA POLYMERASE, DNA REPLICATION, NUCLEOTIDYL 2 TRANSFERASE, SEQUENCING, THIOREDOXIN, PROCESSIVITY FACTOR, 3 COMPLEX (HYDROLASE/ELECTRON TRANSPORT/DNA)
599	1t7p	B	42	132	8e-27	0.53	0.63		DNA POLYMERASE; CHAIN: A; THIOREDOXIN; CHAIN: B; DNA; CHAIN: P, T;	T7 DNA POLYMERASE, DNA REPLICATION, NUCLEOTIDYL 2 TRANSFERASE, SEQUENCING, THIOREDOXIN, PROCESSIVITY FACTOR, 3 COMPLEX (HYDROLASE/ELECTRON TRANSPORT/DNA)
599	1thx		153	264	1.1e-21			55.58	THIOREDOXIN; 1THX 5 CHAIN: NULL; 1THX 6	ELECTRON TRANSPORT THIOREDOXIN 2; 1THX 7 OXIDO-REDUCTASE 1THX 16
599	1thx		158	263	1.1e-21	0.87	1.00		THIOREDOXIN; 1THX 5 CHAIN: NULL; 1THX 6	ELECTRON TRANSPORT THIOREDOXIN 2; 1THX 7 OXIDO-REDUCTASE 1THX 16
599	1thx		2	39	9e-06	-0.23	0.05		THIOREDOXIN; 1THX 5 CHAIN: NULL; 1THX 6	ELECTRON TRANSPORT THIOREDOXIN 2; 1THX 7 OXIDO-REDUCTASE 1THX 16
599	1tof		153	265	1.6e-23			61.47	THIOREDOXIN H; CHAIN: NULL;	ELECTRON TRANSPORT HTRX, HCH1, CH1; OXIDOREDUCTASE, ELECTRON TRANSPORT

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
599	1tof		161	262	1.6e-23	0.81	0.94		THIOREDOXIN H; CHAIN: NULL;	ELECTRON TRANSPORT HTRX, HCH1, CH1; OXIDOREDUCTASE, ELECTRON TRANSPORT
599	1tof		40	135	3.2e-24	-0.19	0.03		THIOREDOXIN H; CHAIN: NULL;	ELECTRON TRANSPORT HTRX, HCH1, CH1; OXIDOREDUCTASE, ELECTRON TRANSPORT
599	2trx	A	153	264	3.2e-29			66.68	ELECTRON TRANSPORT THIOREDOXIN 2TRXA 2 2TRXA 3	
599	2trx	A	158	259	3.2e-29	0.53	1.00		ELECTRON TRANSPORT THIOREDOXIN 2TRXA 2 2TRXA 3	
599	2trx	A	39	132	3.2e-27	0.67	0.96		ELECTRON TRANSPORT THIOREDOXIN 2TRXA 2 2TRXA 3	
602	1bgl	A	79	201	6.4e-07	-0.30	0.05		STAT3B; CHAIN: A; 18-MER DESOXYOLIGONUCLEOTIDE; CHAIN: B;	COMPLEX (TRANSCRIPTION FACTOR/DNA) TRANSCRIPTION FACTOR, PROTEIN-DNA COMPLEX, CYTOKINE 2 ACTIVATION, COMPLEX (TRANSCRIPTION FACTOR/DNA)
602	1ez3	A	65	216	1.6e-08	-0.63	0.09		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
604	1b8q	A	8	127	4.8e-15	0.42	0.29		NEURONAL NITRIC OXIDE SYNTHASE; CHAIN: A; HEPTAPEPTIDE; CHAIN: B;	OXIDOREDUCTASE PDZ DOMAIN, NNOS, NITRIC OXIDE SYNTHASE
604	1be9	A	1	121	9.6e-25			71.82	PSD-95; CHAIN: A; CRIPT; CHAIN: B;	PEPTIDE RECOGNITION PEPTIDE RECOGNITION, PROTEIN LOCALIZATION
604	1be9	A	3	104	9.6e-25	0.62	1.00		PSD-95; CHAIN: A; CRIPT; CHAIN: B;	PEPTIDE RECOGNITION PEPTIDE RECOGNITION, PROTEIN LOCALIZATION
604	1pdr		10	104	3.2e-22	0.80	1.00		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
604	1pdr		8	108	3.2e-22			66.26	HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
604	1qau	A	11	123	1.6e-13	0.79	0.87		NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: A;	OXIDOREDUCTASE BETA-FINGER

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
604	1qav	A	10	97	1.1e-20	1.10	1.00		ALPHA-1 SYNTROPHIN (RESIDUES 77-171); CHAIN: A; NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: B;	MEMBRANE PROTEIN/OXIDOREDUCTASE BETA-FINGER, HETERODIMER
604	1qlc	A	13	99	1.8e-24	1.16	1.00		POSTSYNAPTIC DENSITY PROTEIN 95; CHAIN: A;	PEPTIDE RECOGNITION PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE SYNTHASE, NMDA RECEPTOR 2 BINDING
604	1qlc	A	9	99	4.8e-24	1.28	1.00		POSTSYNAPTIC DENSITY PROTEIN 95; CHAIN: A;	PEPTIDE RECOGNITION PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE SYNTHASE, NMDA RECEPTOR 2 BINDING
604	3pdz	A	13	102	3.2e-22	0.91	1.00		TYROSINE PHOSPHATASE (PTP-BAS, TYPE 1); CHAIN: A;	HYDROLASE PDZ DOMAIN, HUMAN PHOSPHATASE, HPTP1E, PTP-BAS, SPECIFICITY 2 OF BINDING
604	3pdz	A	13	99	1.8e-22	1.22	1.00		TYROSINE PHOSPHATASE (PTP-BAS, TYPE 1); CHAIN: A;	HYDROLASE PDZ DOMAIN, HUMAN PHOSPHATASE, HPTP1E, PTP-BAS, SPECIFICITY 2 OF BINDING
608	1b8i	A	43	93	0.0072	-0.52	0.23		ULTRABITHORAX HOMEOTIC PROTEIN IV; CHAIN: A; HOMEOBOX PROTEIN EXTRADENTICLE; CHAIN: B; DNA (5'- CHAIN: C; DNA (5'- CHAIN: D;	TRANSCRIPTION/DNA ULTRABITHORAX; PBX PROTEIN; DNA BINDING, HOMEODOMAIN, HOMEOTIC PROTEINS, DEVELOPMENT, 2 SPECIFICITY
609	1pft		5	31	0.0085	-0.56	0.16		TFIIB; CHAIN: NULL;	TRANSCRIPTION INITIATION PFTFIIBN; N-TERMINAL DOMAIN, TFIIB, TRANSCRIPTION INITIATION FACTOR
611	1c83	A	2	146	6.4e-29	-0.22	0.07		PROTEIN-TYROSINE PHOSPHATASE 1B; CHAIN: A;	HYDROLASE PTP1B; HYDROLASE, PHOSPHORYLATION, LIGAND, INHIBITOR
611	1gwz		2	148	1.6e-31	-0.59	0.00		SHP-1; CHAIN: NULL;	HYDROLASE PROTEIN-TYROSINE PHOSPHATASE; HYDROLASE, PROTEIN TYROSINE PHOSPHATASE, CATALYTIC DOMAIN, 2 WPD LOOP, SH2 DOMAIN
611	1lar	A	3	147	3.2e-35	0.02	0.04		LAR; CHAIN: A, B;	HYDROLASE TYROSINE PHOSPHATEASE, LAR PROTEIN
611	1mkp		27	149	1.6e-24	0.30	0.99		PYST1; CHAIN: NULL;	HYDROLASE DUAL SPECIFICITY

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										PHOSPHATASE, MAP KINASE HYDROLASE
611	1mkp		6	149	1.6e-24			56.76	PYST1; CHAIN: NULL;	HYDROLASE DUAL SPECIFICITY PHOSPHATASE, MAP KINASE HYDROLASE
611	1rpm	A	1	147	4.8e-34	0.25	0.09		RECEPTOR PROTEIN TYROSINE PHOSPHATASE MU; CHAIN: A, B;	RECEPTOR D1; RECEPTOR, PHOSPHATASE, SIGNAL TRANSDUCTION, ADHESION, 2 HYDROLASE
611	1vhr	A	3	138	3.6e-19	0.75	0.96		HUMAN VH1-RELATED DUAL-SPECIFICITY PHOSPHATASE CHAIN: A, B;	HYDROLASE VHR; HYDROLASE, PROTEIN DUAL-SPECIFICITY PHOSPHATASE
611	1yfo	A	5	147	1.6e-29	-0.20	0.69		RECEPTOR PROTEIN TYROSINE PHOSPHATASE ALPHA; CHAIN: A, B;	HYDROLASE D1; HYDROLASE, SIGNAL TRANSDUCTION, RECEPTOR, GLYCOPROTEIN, 2 PHOSPHORYLATION, SIGNAL
611	2shp	A	2	147	1.1e-32	-0.22	0.19		SHP-2; CHAIN: A, B;	TYROSINE PHOSPHATASE SYP, SHPTP-2; TYROSINE PHOSPHATASE, INSULIN SIGNALING, SH2 PROTEIN
612	1adq	L	4	188	6.4e-17	-0.27	0.18		IGG4 REA; CHAIN: A; RF-AN IGM/LAMBDA; CHAIN: H, L;	COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN) COMPLEX (IMMUNOGLOBULIN/AUTOANTIGEN), RHEUMATOID FACTOR 2 AUTO-ANTIBODY COMPLEX
612	1bih	A	1	380	1.6e-51			97.76	HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
612	1igt	B	1	380	3.2e-20			85.83	IGG2A INTACT ANTIBODY - MAB231; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN V REGION C REGION, IMMUNOGLOBULIN
612	1igy	B	1	379	6.4e-21			82.67	IGG1 INTACT ANTIBODY MAB61.1.3; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN, V REGION, C REGION, HINGE REGION
612	1kb5	L	195	364	6.4e-13	0.00	-0.18		KB5-C20 T-CELL ANTIGEN RECEPTOR; CHAIN: A, B; ANTIBODY DESIRE-1; CHAIN: L, H;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) TCR VAPLHA VBETA DOMAIN; T-CELL RECEPTOR, STRAND SWITCH, FAB, ANTICLONOTYPIC, 2

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										(IMMUNOGLOBULIN/RECEPTOR)
612	1mco	H	1	378	3.2e-24			87.85	IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION 1MCO 3	
612	8fab	A	5	188	3.2e-17	-0.29	0.21		IMMUNOGLOBULIN FAB FRAGMENT FROM HUMAN IMMUNOGLOBULIN IGG1 (LAMBDA, HIL) 8FAB 3	
613	1a3k		1	125	6.4e-40			66.61	GALECTIN-3; CHAIN: NULL;	GALECTIN GALECTIN, GALAPTIN, LECTIN, IGE-BINDING PROTEIN
613	1a3k		3	123	6.4e-40	0.65	1.00		GALECTIN-3; CHAIN: NULL;	GALECTIN GALECTIN, GALAPTIN, LECTIN, IGE-BINDING PROTEIN
613	1a78	A	1	117	3.2e-24	0.73	0.69		GALECTIN-1; CHAIN: A, B;	LECTIN S-LECTIN GALECTIN; S-LECTIN, CARBOHYDRATE BINDING, COMPLEX (LECTIN/SACCHARIDE)
613	1bkz	A	1	125	9.6e-38			73.96	GALECTIN-7; CHAIN: A, B;	LECTIN GALAPTIN, LECTIN, GALECTIN, CARBOHYDRATE BINDING
613	1bkz	A	2	124	9.6e-38	0.67	1.00		GALECTIN-7; CHAIN: A, B;	LECTIN GALAPTIN, LECTIN, GALECTIN, CARBOHYDRATE BINDING
613	1cll	A	1	122	3.6e-23	0.42	1.00		CONGERIN I; CHAIN: A;	SUGAR BINDING PROTEIN GALECTIN, LECTIN, BETA-GALACTOSE-BINDING, SUGAR BINDING 2 PROTEIN
613	1cll	A	2	124	6.4e-21	0.44	0.99		CONGERIN I; CHAIN: A;	SUGAR BINDING PROTEIN GALECTIN, LECTIN, BETA-GALACTOSE-BINDING, SUGAR BINDING 2 PROTEIN
613	1hlc	A	1	123	1.3e-30	0.62	1.00		LECTIN LECTIN (HUMAN L-14-II) COMPLEXED WITH LACTOSE 1HLC 3	
613	1hlc	A	1	125	1.3e-30			52.58	LECTIN LECTIN (HUMAN L-14-II) COMPLEXED WITH LACTOSE 1HLC 3	
613	1lcl		1	123	8e-32	0.80	1.00		LYSOPHOSPHOLIPASE; CHAIN: NULL;	SERINE ESTERASE CHARCOT-LEYDEN CRYSTAL PROTEIN; CHARCOT-LEYDEN CRYSTAL PROTEIN, SERINE ESTERASE
613	1qmj	A	1	124	3.2e-29	0.45	1.00		BETA-GALACTOSIDE-BINDING LECTIN; CHAIN: A, B;	GALECTIN 16 KD LECTIN, C-16 GALECTIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
613	1slt	A	1	124	1.4e-33	0.40	1.00		LECTIN S-LECTIN (A VERTEBRATE 14 KDA BETA-GALACTOSIDE BINDING 1SLT 3 PROTEIN) COMPLEX WITH N-ACETYLLACTOSAMINE 1SLT 4	
614	1a2y	A	1024	1102	0.0054	-0.07	0.13		MONOCLONAL ANTIBODY D1.3; CHAIN: A, B; LYSOZYME; CHAIN: C;	COMPLEX (IMMUNOGLOBULIN/HYDROLASE) COMPLEX (IMMUNOGLOBULIN/HYDROLASE), IMMUNOGLOBULIN V 2 REGION, SIGNAL, HYDROLASE, GLYCOSIDASE, BACTERIOLYTIC 3 ENZYME, EGG WHITE
614	1a7q	L	1024	1102	0.0036	0.17	0.04		MONOCLONAL ANTIBODY D1.3; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, VARIANT
614	1ar1	D	1024	1102	0.0072	0.38	0.21		CYTOCHROME C OXIDASE; CHAIN: A, B; ANTIBODY FV FRAGMENT; CHAIN: C, D;	COMPLEX (OXIDOREDUCTASE/ANTIBODY) CYTOCHROME AA3, COMPLEX IV, FERROCYTOCHROME C, COMPLEX (OXIDOREDUCTASE/ANTIBODY), ELECTRON TRANSPORT, 2 TRANSMEMBRANE, CYTOCHROME OXIDASE, ANTIBODY COMPLEX
614	1bvk	A	1024	1102	0.0072	0.49	0.13		HULYS11; CHAIN: A, B, D, E; LYSOZYME; CHAIN: C, F;	COMPLEX (HUMANIZED ANTIBODY/HYDROLASE) MURAMIDASE; HUMANIZED ANTIBODY, ANTIBODY COMPLEX, FV, ANTI-LYSOZYME, 2 COMPLEX (HUMANIZED ANTIBODY/HYDROLASE)
614	1ehd	A	430	462	0.0014	2.14	0.05		AGGLUTININ ISOLECTIN VI; CHAIN: A	PLANT PROTEIN TWO HOMOLOGOUS HEVEIN-LIKE DOMAINS
614	1eis	A	430	463	0.0036	2.14	0.03		AGGLUTININ ISOLECTIN VI/AGGLUTININ ISOLECTIN V; CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTIGEN
614	1en2	A	430	463	0.0018	1.47	0.12		AGGLUTININ ISOLECTIN I/AGGLUTININ ISOLECTIN V/ CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTIGEN, SACCHARIDE BINDING
614	1jhl	L	1024	1102	0.0036	0.28	0.19		COMPLEX(ANTIBODY-ANTIGEN)	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									FV FRAGMENT (IGG1, KAPPA) (LIGHT AND HEAVY VARIABLE DOMAINS 1JHL 3 NON-COVALENTLY ASSOCIATED) OF MONOCLONAL ANTI-HEN EGG 1JHL 4 LYSOZYME ANTIBODY D11.15 COMPLEX WITH PHEASANT EGG 1JHL 5 LYSOZYME 1JHL 6	
614	1jrh	L	1024	1108	0.0072	0.28	0.13		ANTIBODY A6; CHAIN: L, H; INTERFERON-GAMMA RECEPTOR ALPHA CHAIN; CHAIN: I;	COMPLEX (ANTIBODY/ANTIGEN) CYTOKINE RECEPTOR, COMPLEX (ANTIBODY/ANTIGEN), 2 TRANSMEMBRANE, GLYCOPROTEIN
614	1osm	A	1268	1486	1.3e-11	0.52	-0.19		OMPK36; CHAIN: A, B, C;	OUTER MEMBRANE PROTEIN OSMOPORIN; OUTER MEMBRANE PROTEIN, NON-SPECIFIC PORIN, OSMOPORIN, 2 BETA-BARREL, TRANSMEMBRANE
614	1osm	A	821	1045	7.2e-12	0.62	-0.20		OMPK36; CHAIN: A, B, C;	OUTER MEMBRANE PROTEIN OSMOPORIN; OUTER MEMBRANE PROTEIN, NON-SPECIFIC PORIN, OSMOPORIN, 2 BETA-BARREL, TRANSMEMBRANE
614	1pho		1291	1474	1.8e-12	0.64	-0.20		OUTER MEMBRANE PROTEIN PHOSPHOPORIN (PHOE) 1PHO 3	
614	1wtl	A	1024	1102	0.009	0.31	0.78		IMMUNOGLOBULIN WAT, A VARIABLE DOMAIN FROM IMMUNOGLOBULIN LIGHT-CHAIN 1WTL 3 (BENCE-JONES PROTEIN) 1WTL 4	
614	2omf		1290	1486	9e-10	0.70	-0.19		MATRIX PORIN OUTER MEMBRANE PROTEIN F; 2OMF 5 CHAIN: NULL; 2OMF 6	INTEGRAL MEMBRANE PROTEIN PORIN MATRIX PORIN, OMPF PORIN; 2OMF 7 PORIN, MEMBRANE PROTEIN 2OMF 12
616	1alh	A	316	372	1.1e-15	0.11	-0.20		QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
616	1mey	G	316	342	4.8e-11	0.28	-0.20		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
616	1tf3	A	316	368	3.2e-14	0.02	-0.20		TRANSCRIPTION FACTOR IIIA; CHAIN: A; 5S RNA GENE; CHAIN: E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) TFIIIA; 5S GENE; NMR, TFIIIA, PROTEIN, DNA, TRANSCRIPTION FACTOR, 5S RNA 2 GENE, DNA BINDING PROTEIN, ZINC FINGER, COMPLEX 3 (TRANSCRIPTION REGULATION/DNA)
618	1av1	A	1	202	3.2e-07			50.15	APOLIPOPROTEIN A-I; CHAIN: A, B, C, D;	LIPID TRANSPORT APO A-I; LIPOPROTEIN, LIPID TRANSPORT, CHOLESTEROL METABOLISM, 2 ATHEROSCLEROSIS, HDL, LCAT-ACTIVATION
618	1cun	A	1	201	0.0014			58.88	ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
628	1a1h	A	391	473	6e-36			77.55	QGSR ZINC FINGER PEPTIDE; CHAIN: A; DUPLEX OLIGONUCLEOTIDE BINDING SITE; CHAIN: B, C;	COMPLEX (ZINC FINGER/DNA) COMPLEX (ZINC FINGER/DNA), ZINC FINGER, DNA-BINDING PROTEIN
628	1mey	C	306	387	2e-45	0.47	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
628	1mey	C	334	415	4e-44	0.59	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
628	1mey	C	334	416	2e-45			97.73	DNA; CHAIN: A, B, D, E;	COMPLEX (ZINC FINGER/DNA) ZINC

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
628	1mey	C	362	443	2e-42	0.06	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
628	1mey	C	390	471	6e-38	0.17	1.00		DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
628	1ubd	C	281	387	4e-43	0.20	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
628	1ubd	C	332	443	6e-52	0.37	1.00		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
628	1ubd	C	332	444	6e-52			85.61	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
628	1ubd	C	360	471	2e-48	0.31	0.99		YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: A, B;	ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
628	2gli	A	273	417	6e-67			97.26	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
628	2gli	A	280	417	1.4e-54	-0.11	0.99		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
628	2gli	A	306	445	6e-67	0.53	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
628	2gli	A	334	471	1.4e-62	0.20	1.00		ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
632	1dd5	A	85	261	4e-42	0.69	1.00		RIBOSOME RECYCLING FACTOR; CHAIN: A;	RIBOSOME THREE-HELIX BUNDLE, BETA-ALPHA-BETA SANDWICH, RIBOSOME
632	1eh1	A	85	262	4e-39	0.41	1.00		RIBOSOME RECYCLING FACTOR; CHAIN: A;	RIBOSOME TRANSLATION, RIBOSOME, HINGE VARIABILITY
639	1apm	E	1	328	6e-41			97.88	TRANSFERASE(PHOSPHOTRANS FERASE) $C-/AMP$-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (C/APK) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (/S139A$) COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	
639	1aq1		13	309	2e-38			94.94	CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE, PHOSPHORYLATION, STAUROSPORINE, 2 CELL DIVISION, MITOSIS, INHIBITION
639	1b6c	B	1	268	1e-65			98.18	FK506-BINDING PROTEIN;	COMPLEX (ISOMERASE/PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: A, C, E, G; TGF-B SUPERFAMILY RECEPTOR TYPE I; CHAIN: B, D, F, H;	KINASE) FKBP12; SERINE/THREONINE-PROTEIN KINASE RECEPTOR R4; COMPLEX (ISOMERASE/PROTEIN KINASE), RECEPTOR 2 SERINE/THREONINE KINASE
639	1bi8	A	14	296	2e-41			96.84	CYCLIN-DEPENDENT KINASE 6; CHAIN: A, C; CYCLIN-DEPENDENT KINASE INHIBITOR; CHAIN: B, D;	COMPLEX (KINASE/INHIBITOR) CDK6; P19INK4D; CYCLIN DEPENDENT KINASE, CYCLIN DEPENDENT KINASE INHIBITORY 2 PROTEIN, CDK, INK4, CELL CYCLE, COMPLEX (KINASE/INHIBITOR) HEADER HELIX
639	1blx	A	1	300	2e-40			94.81	CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
639	1byg	A	10	268	8e-74			142.73	C-TERMINAL SRC KINASE; CHAIN: A;	TRANSFERASE CSK; PROTEIN KINASE, C-TERMINAL SRC KINASE, PHOSPHORYLATION, 2 STAUROSPORINE, TRANSFERASE
639	1cmk	E	1	328	6e-41			92.82	PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	
639	1ctp	E	1	320	1e-40			95.40	TRANSFERASE(PHOSPHOTRANS FERASE) CAMP-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (CAPK) 1CTP 3 (CATALYTIC SUBUNIT) 1CTP 4	
639	1fgk	A	2	268	8e-72			146.59	FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
639	1fgk	B	1	267	4e-71			146.30	FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										PROTEIN KINASE, ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
639	1hcl		13	309	2e-41			99.30	HUMAN CYCLIN-DEPENDENT KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; TRANSFERASE, SERINE/THREONINE PROTEIN KINASE, ATP-BINDING, 2 CELL CYCLE, CELL DIVISION, MITOSIS, PHOSPHORYLATION
639	1ir3	A	1	282	4e-74			147.38	INSULIN RECEPTOR; CHAIN: A; PEPTIDE SUBSTRATE; CHAIN: B;	COMPLEX (TRANSFERASE/SUBSTRATE) TYROSINE KINASE, SIGNAL TRANSDUCTION, PHOSPHOTRANSFERASE, 2 COMPLEX (KINASE/PEPTIDE SUBSTRATE/ATP ANALOG), ENZYME, 3 COMPLEX (TRANSFERASE/SUBSTRATE)
639	1koa		1	420	8e-45			110.26	TWITCHIN; CHAIN: NULL;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
639	1kob	A	1	328	2e-41			97.87	TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
647	1bu7	A	25	469	7.2e-63			119.02	CYTOCHROME P450; CHAIN: A, B;	OXIDOREDUCTASE FATTY ACID HYDROXYLASE; FATTY ACID MONOOXYGENASE, HEMOPROTEIN, P450 REMARK
647	1cpt		5	468	7.2e-25			76.99	OXIDOREDUCTASE(OXYGENASE) CYTOCHROME P450-TERP 1CPT 3	
647	1oxa		3	467	7.2e-33			82.30	CYTOCHROME P450 ERYF; 1OXA 5 CHAIN: NULL 1OXA 6	OXIDOREDUCTASE (OXYGENASE)
657	1mey	C	430	512	3.6e-50			96.26	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
657	1tf6	A	122	288	7.2e-38			100.65	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										INITIATION, ZINC FINGER PROTEIN
657	1ubd	C	264	372	1.4e-55			80.97	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
657	2gli	A	94	233	1.8e-69			90.74	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
659	1av1	A	1	199	0.00036			53.84	APOLIPOPROTEIN A-I; CHAIN: A, B, C, D;	LIPID TRANSPORT APO A-I; LIPOPROTEIN, LIPID TRANSPORT, CHOLESTEROL METABOLISM, 2 ATHEROSCLEROSIS, HDL, LCAT-ACTIVATION
659	1cun	A	46	196	0.0016	-0.09	0.19		ALPHA SPECTRIN; CHAIN: A, B, C;	STRUCTURAL PROTEIN TWO REPEATS OF SPECTRIN, ALPHA HELICAL LINKER REGION, 2 2 TANDEM 3-HELIX COILED-COILS, STRUCTURAL PROTEIN
659	1fxk	A	75	165	0.00011	0.07	0.03		PREFOLDIN; CHAIN: A; PREFOLDIN; CHAIN: B; PREFOLDIN; CHAIN: C;	CHAPERONE ARCHAEAL PROTEIN
660	1b7t	A	8	176	3.6e-63	0.19	1.00		MYOSIN HEAVY CHAIN; CHAIN: A; MYOSIN REGULATORY LIGHT CHAIN; CHAIN: Y; MYOSIN ESSENTIAL LIGHT CHAIN; CHAIN: Z;	MYOSIN MYOSIN MOTOR
660	1br1	A	10	176	1.1e-63	-0.04	1.00		MYOSIN; CHAIN: A, B, C, D, E, F, G, H;	MUSCLE PROTEIN MDE; MUSCLE PROTEIN
660	1dfk	A	8	176	7.2e-63	0.08	1.00		MYOSIN HEAD; CHAIN: A; MYOSIN HEAD; CHAIN: Y; MYOSIN HEAD; CHAIN: Z;	CONTRACTILE PROTEIN MYOSIN MOTOR, CONFORMATIONAL CHANGES
660	1lvk	.	12	176	3.6e-54	0.03	1.00		MYOSIN; CHAIN: NULL;	CONTRACTILE PROTEIN MYOSIN, DICTYOSTELIUM, MOTOR, MANT, ATPASE, ACTIN-BINDING, 2 COILED COIL

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
660	1lvk		5	176	1.4e-58	0.04	1.00		MYOSIN; CHAIN: NULL;	CONTRACTILE PROTEIN MYOSIN, DICTYOSTELIUM, MOTOR, MANT, ATPASE, ACTIN-BINDING, 2 COILED COIL
660	1mnd		12	176	3.6e-52	0.16	0.99		MYOSIN; CHAIN: NULL;	CONTRACTILE PROTEIN ATPASE, MYOSIN, COILED COIL, ACTIN-BINDING, ATP-BINDING, 2 HEPTAD REPEAT PATTERN, METHYLATION, ALKYLATION, 3 PHOSPHORYLATION, CONTRACTILE PROTEIN
660	1mnd		18	176	2e-56	-0.01	1.00		MYOSIN; CHAIN: NULL;	CONTRACTILE PROTEIN ATPASE, MYOSIN, COILED COIL, ACTIN-BINDING, ATP-BINDING, 2 HEPTAD REPEAT PATTERN, METHYLATION, ALKYLATION, 3 PHOSPHORYLATION, CONTRACTILE PROTEIN
660	2mys	A	3	176	3.6e-57	-0.02	1.00		MYOSIN; CHAIN: A, B, C;	MUSCLE PROTEIN MUSCLE PROTEIN, MYOSIN SUBFRAGMENT-1, MYOSIN HEAD, 2 MOTOR PROTEIN
685	1hme		153	224	5.4e-16	0.18	0.84		DNA-BINDING HIGH MOBILITY GROUP PROTEIN FRAGMENT-B (HMGB) (DNA-BINDING 1HME 3 HMG-BOX DOMAIN B OF RAT HMG1) (NMR, 1 STRUCTURE) 1HME 4	
685	1hry	A	152	224	1.8e-16			55.78	HUMAN SRY; 1HRY 6 CHAIN: A; 1HRY 7 DNA; 1HRY 9 CHAIN: B; 1HRY 10	COMPLEX (DNA-BINDING PROTEIN/DNA)
685	1hry	A	155	224	1.8e-16	-0.14	0.35		HUMAN SRY; 1HRY 6 CHAIN: A; 1HRY 7 DNA; 1HRY 9 CHAIN: B; 1HRY 10	COMPLEX (DNA-BINDING PROTEIN/DNA)
685	1hsm		153	228	5.4e-17	0.15	0.60		DNA-BINDING HIGH MOBILITY GROUP PROTEIN 1 (HMG1) BOX 2, COMPLEXED WITH 1HSM 3 MERCAPTOETHANOL (NMR, MINIMIZED AVERAGE STRUCTURE) 1HSM 4	
685	2lef	A	152	237	1.4e-21			130.26	LYMPHOID ENHANCER-BINDING	GENE REGULATION/DNA LEF-1 HMG;

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									FACTOR; CHAIN: A; DNA (5'- CHAIN: B; DNA (5'- CHAIN: C;	LEF1, HMG, TCR-A, TRANSCRIPTION FACTOR, DNA BINDING, DNA 2 BENDING, COMPLEX (HMG DOMAIN/DNA), GENE REGULATION/DNA
685	2lef	A	153	237	1.8e-13	0.47	1.00		LYMPHOID ENHANCER-BINDING FACTOR; CHAIN: A; DNA (5'- CHAIN: B; DNA (5'- CHAIN: C;	GENE REGULATION/DNA LEF-1 HMG; LEF1, HMG, TCR-A, TRANSCRIPTION FACTOR, DNA BINDING, DNA 2 BENDING, COMPLEX (HMG DOMAIN/DNA), GENE REGULATION/DNA
685	2lef	A	154	228	1.4e-21	0.60	1.00		LYMPHOID ENHANCER-BINDING FACTOR; CHAIN: A; DNA (5'- CHAIN: B; DNA (5'- CHAIN: C;	GENE REGULATION/DNA LEF-1 HMG; LEF1, HMG, TCR-A, TRANSCRIPTION FACTOR, DNA BINDING, DNA 2 BENDING, COMPLEX (HMG DOMAIN/DNA), GENE REGULATION/DNA
689	1a9n	B	139	219	8e-23	0.61	0.98		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
689	1b7f	A	1	109	6e-27	0.74	1.00		SXL-LETHAL PROTEIN; CHAIN: A, B; RNA (5'- R(P*GP*UP*UP*GP*UP*UP*UP*U P*UP*UP*UP*U)- CHAIN: P, Q;	RNA-BINDING PROTEIN/RNA TRA PRE-MRNA; SPLICING REGULATION, RNP DOMAIN, RNA COMPLEX
689	1b7f	A	32	211	1.8e-37	0.92	1.00		SXL-LETHAL PROTEIN; CHAIN: A, B; RNA (5'- R(P*GP*UP*UP*GP*UP*UP*UP*U P*UP*UP*UP*U)- CHAIN: P, Q;	RNA-BINDING PROTEIN/RNA TRA PRE-MRNA; SPLICING REGULATION, RNP DOMAIN, RNA COMPLEX
689	1b7f	A	32	211	6e-47			91.55	SXL-LETHAL PROTEIN; CHAIN: A, B; RNA (5'- R(P*GP*UP*UP*GP*UP*UP*UP*U P*UP*UP*UP*U)- CHAIN: P, Q;	RNA-BINDING PROTEIN/RNA TRA PRE-MRNA; SPLICING REGULATION, RNP DOMAIN, RNA COMPLEX
689	1b7f	A	35	211	6e-47	0.96	1.00		SXL-LETHAL PROTEIN; CHAIN: A, B; RNA (5'- R(P*GP*UP*UP*GP*UP*UP*UP*U P*UP*UP*UP*U)- CHAIN: P, Q;	RNA-BINDING PROTEIN/RNA TRA PRE-MRNA; SPLICING REGULATION, RNP DOMAIN, RNA COMPLEX
689	1cvj	A	1	115	4e-28	0.57	1.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'- R(*AP*AP*AP*AP*AP*AP*AP*AP	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	
689	1cvj	A	34	217	2e-48			107.38	POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP *AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
689	1cvj	A	35	217	2e-48	1.03	1.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP *AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
689	1cvj	A	36	217	5.4e-37	0.97	1.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP *AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
689	1cvj	B	34	202	8e-42			93.53	POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP *AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
689	1cvj	B	35	198	8e-42	1.03	1.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP *AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
689	1cvj	B	36	197	3.6e-31	1.06	1.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP *AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
689	1cvj	F	35	189	2e-33	0.71	1.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
689	1cvj	H	34	193	6e-33			63.68	POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
689	1cvj	H	35	189	6e-33	0.84	1.00		POLYDENYLATE BINDING PROTEIN 1; CHAIN: A, B, C, D, E, F, G, H; RNA (5'-R(*AP*AP*AP*AP*AP*AP*AP*AP*AP*AP*A)-3'); CHAIN: M, N, O, P, Q, R, S, T;	GENE REGULATION/RNA POLY(A) BINDING PROTEIN 1, PABP 1; RRM, PROTEIN-RNA COMPLEX, GENE REGULATION/RNA
689	1cx0	A	138	221	6e-25	0.90	1.00		U1A PROTEIN; CHAIN: A; HDV RIBOZYME SELF-CLEAVED; CHAIN: B;	RNA BINDING PROTEIN/RNA NESTED DOUBLE PSEUDOKNOT RNA STRUCTURE
689	1d8z	A	138	217	4e-24	0.71	1.00		HU ANTIGEN C; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
689	1d8z	A	35	112	1.4e-26	1.19	1.00		HU ANTIGEN C; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
689	1d9a	A	138	216	1.8e-24	0.97	1.00		HU ANTIGEN C; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
689	1fht		138	225	2e-23	0.60	1.00		U1 SMALL NUCLEAR RIBONUCLEOPROTEIN A; CHAIN: NULL;	RIBONUCLEOPROTEIN U1A117; RIBONUCLEOPROTEIN, RNP DOMAIN, SPLICEOSOME
689	1fj7	A	138	218	6e-24	0.24	0.09		NUCLEOLIN RBD1; CHAIN: A;	STRUCTURAL PROTEIN PROTEIN C23; RNP, RBD, RRM, RNA BINDING DOMAIN, NUCLEOLUS
689	1fj7	A	18	116	2e-27	0.32	0.75		NUCLEOLIN RBD1; CHAIN: A;	STRUCTURAL PROTEIN PROTEIN C23; RNP, RBD, RRM, RNA BINDING DOMAIN, NUCLEOLUS
689	1fjc	A	138	217	1e-21	0.38	0.95		NUCLEOLIN RBD2; CHAIN: A;	STRUCTURAL PROTEIN PROTEIN C23; RNP, RBD, RRM, RNA BINDING DOMAIN,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										NUCLEOLUS
689	1hal		29	207	1.3e-44			74.26	HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
689	1hal		29	211	1.3e-44	0.53	1.00		HNRNP A1; CHAIN: NULL;	NUCLEAR PROTEIN HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1, NUCLEAR PROTEIN, HNRNP, RBD, RRM, RNP, RNA BINDING, 2 RIBONUCLEOPROTEIN
689	1hdl	A	139	211	2e-21	0.96	1.00		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN D0; CHAIN: A;	RNA BINDING PROTEIN RNA-BINDING DOMAIN
689	1nrc	A	138	213	1e-21	0.01	0.99		RIBONUCLEOPROTEIN PROTEIN FROM U1 SMALL NUCLEAR RIBONUCLEOPROTEIN (SNRNP U1) 1NRC 3 (N-TERMINAL FRAGMENT, RESIDUES 1 - 95) MUTANT WITH GLN 85 1NRC 4 REPLACED BY CYS (Q85C) 1NRC 5	
689	1qm9	A	138	251	4e-19	0.05	-0.01		POLYPYRIMIDINE TRACT-BINDING PROTEIN; CHAIN: A;	RIBONUCLEOPROTEIN PTB, PTB-C198, HETEROGENEOUS NUCLEAR POLYPYRIMIDINE TRACT BINDING PROTEIN, RNP, RNA, SPICING, 2 TRANSLATION
689	1qm9	A	35	212	4e-44	0.27	0.87		POLYPYRIMIDINE TRACT-BINDING PROTEIN; CHAIN: A;	RIBONUCLEOPROTEIN PTB, PTB-C198, HETEROGENEOUS NUCLEAR POLYPYRIMIDINE TRACT BINDING PROTEIN, RNP, RNA, SPICING, 2 TRANSLATION
689	2sxl		138	217	2e-25	0.91	1.00		SEX-LETHAL PROTEIN; CHAIN: NULL;	RNA-BINDING DOMAIN RNA-BINDING DOMAIN, ALTERNATIVE SPLICING
689	2sxl		35	115	4e-28	1.34	1.00		SEX-LETHAL PROTEIN; CHAIN: NULL;	RNA-BINDING DOMAIN RNA-BINDING DOMAIN, ALTERNATIVE SPLICING
689	2up1	A	1	116	6e-29	0.63	1.00		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP A1, UP1; COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA; CHAIN: B;	(RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
689	2up1	A	28	217	1.8e-47	0.51	1.00		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1; CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA; CHAIN: B;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP A1, UP1; COMPLEX (RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
689	2up1	A	28	219	8e-50			79.48	HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1; CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA; CHAIN: B;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP A1, UP1; COMPLEX (RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
689	2up1	A	35	218	8e-50	0.93	1.00		HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN A1; CHAIN: A; 12-NUCLEOTIDE SINGLE-STRANDED TELOMETRIC DNA; CHAIN: B;	COMPLEX (RIBONUCLEOPROTEIN/DNA) HNRNP A1, UP1; COMPLEX (RIBONUCLEOPROTEIN/DNA), HETEROGENEOUS NUCLEAR 2 RIBONUCLEOPROTEIN A1
689	3sxl	A	33	201	1.8e-35	0.67	1.00		SEX-LETHAL; CHAIN: A, B, C;	RNA BINDING DOMAIN RNA BINDING DOMAIN, RBD, RNA RECOGNITION MOTIF, RRM, 2 SPLICING INHIBITOR, TRANSLATIONAL INHIBITOR, SEX 3 DETERMINATION, X CHROMOSOME DOSAGE COMPENSATION
689	3sxl	A	34	204	2e-43			78.56	SEX-LETHAL; CHAIN: A, B, C;	RNA BINDING DOMAIN RNA BINDING DOMAIN, RBD, RNA RECOGNITION MOTIF, RRM, 2 SPLICING INHIBITOR, TRANSLATIONAL INHIBITOR, SEX 3 DETERMINATION, X CHROMOSOME DOSAGE COMPENSATION
689	3sxl	A	35	204	2e-43	0.93	1.00		SEX-LETHAL; CHAIN: A, B, C;	RNA BINDING DOMAIN RNA BINDING DOMAIN, RBD, RNA RECOGNITION MOTIF, RRM, 2 SPLICING INHIBITOR, TRANSLATIONAL INHIBITOR, SEX 3 DETERMINATION, X CHROMOSOME DOSAGE COMPENSATION
692	1a09	A	156	236	4e-05	0.12	0.63		C-SRC TYROSINE KINASE;	COMPLEX (TRANSFERASE/PEPTIDE)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: A, B; ACE-FORMYL PHOSPHOTYR-GLU-(N,N-DIPENTYL AMINE); CHAIN: C, D;	COMPLEX (TRANSFERASE/PEPTIDE)
692	1aot	F	156	236	2e-05	-0.07	0.60		FYN PROTEIN-TYROSINE KINASE; CHAIN: F; PHOSPHOTYROSYL PEPTIDE; CHAIN: P	COMPLEX (PROTO-ONCOGENE/EARLY PROTEIN) SRC HOMOLOGY 2 DOMAIN; SH2 DOMAIN, SIGNAL TRANSDUCTION, PEPTIDE COMPLEX, 2 COMPLEX (PROTO-ONCOGENE/EARLY PROTEIN)
692	1aya	A	156	236	8e-10	0.15	0.84		HYDROLASE(SH2 DOMAIN) TYROSINE PHOSPHATASE SYP (N-TERMINAL SH2 DOMAIN) 1AYA 3 (PTP1D, SHPTP2) (E.C.3.1.3.48) COMPLEXED WITH THE PEPTIDE 1AYA 4 PDGFR-1009 1AYA 5	
692	1bkl		156	236	2e-05	0.18	0.77		PP60 V-SRC TYROSINE KINASE TRANSFORMING PROTEIN; CHAIN: NULL;	V-SRC SH2 DOMAIN SRC SH2; V-SRC SH2 DOMAIN, PHOSPHOTYROSINE RECOGNITION DOMAIN, PP60 2 SRC SH2 DOMAIN
692	1blj		156	236	1.8e-05	0.11	0.28		P55 BLK PROTEIN TYROSINE KINASE; CHAIN: NULL;	PHOSPHORYLATION SIGNAL TRANSDUCTION, TYROSINE KINASE, TRANSFERASE, 2 PHOSPHOTRANSFERASE, PHOSPHORYLATION
692	1btn		23	109	0.0016	-0.14	0.06		BETA-SPECTRIN; 1BTN 4 CHAIN: NULL; 1BTN 5	SIGNAL TRANSDUCTION PROTEIN
692	1cwd	L	156	236	0.00012	0.26	0.69		P56LCK TYROSINE KINASE; CHAIN: L; PHOSPHONOPEPTIDE CHAIN: P;	COMPLEX (PHOSPHOTRANSFERASE/PEPTIDE) PHOSPHOTRANSFERASE, COMPLEX (PHOSPHOTRANSFERASE/PEPTIDE)
692	1fao	A	18	117	4e-06	0.35	0.01		DUAL ADAPTOR OF PHOSPHOTYROSINE AND 3-CHAIN: A;	SIGNALING PROTEIN DAPP1, PHISH, BAM32; PLECKSTRIN, 3-PHOSPHOINOSITIDES, INOSITOL TETRAKISPHOSPHATE 2 SIGNAL TRANSDUCTION PROTEIN, ADAPTOR PROTEIN
692	1fb8	A	23	109	4e-06	0.51	0.64		DUAL ADAPTOR OF PHOSPHOTYROSINE AND 3-	SIGNALING PROTEIN DAPP1, PHISH, BAM32; PLECKSTRIN, 3-

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: A;	PHOSPHOINOSITIDES, INOSITOL TETRAKISPHOSPHATE 2 SIGNAL TRANSDUCTION PROTEIN, ADAPTOR PROTEIN
692	1fhs		156	236	2e-08	0.10	0.93		GROWTH FACTOR RECEPTOR BOUND PROTEIN-2; CHAIN: NULL;	SH2 DOMAIN GRB2; GRB2, SH2 DOMAIN, PROTEIN NMR, SOLUTION STRUCTURES
692	1pls		18	117	1e-06	0.14	0.16		PHOSPHORYLATION PLECKSTRIN (N-TERMINAL PLECKSTRIN HOMOLOGY DOMAIN) MUTANT 1PLS 3 WITH LEU GLU (HIS)6 ADDED TO THE C TERMINUS 1PLS 4 (INS(G105-LEHHHHHH)) (NMR, 25 STRUCTURES) 1PLS 5	
692	1sha	A	156	236	2e-05	0.16	0.96		PHOSPHOTRANSFERASE V-SRC TYROSINE KINASE TRANSFORMING PROTEIN (PHOSPHOTYROSINE 1SHA 3 RECOGNITION DOMAIN SH2) (E.C.2.7.1.112) COMPLEX WITH 1SHA 4 PHOSPHOPEPTIDE A (TYR-VAL-PRO-MET-LEU, PHOSPHORYLATED TYR) 1SHA 5	
692	2pld	A	156	236	1.8e-06	-0.02	0.31		PHOSPHORIC DIESTER HYDROLASE PHOSPHOLIPASE C-GAMMA-1 (E.C.3.1.4.11) (C-TERMINAL SH2 2PLD 3 DOMAIN COMPRISING RESIDUES 663 - 759) COMPLEXED WITH A 2PLD 4 PHOSPHOPEPTIDE FROM THE PLATELET-DERIVED GROWTH FACTOR 2PLD 5 RECEPTOR (RESIDUES 1018 - 1029: ASP-ASN-ASP-PTYR-ILE-ILE- 2PLD 6 PRO-LEU-PRO-ASP-PRO-LYS) (NMR, MINIMIZED AVERAGE STRUCTURE) 2PLD 7	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
692	1a09	A	156	236	4e-05	0.12	0.63		C-SRC TYROSINE KINASE; CHAIN: A, B; ACE-FORMYL PHOSPHOTYR-GLU-(N,N-DIPENTYL AMINE); CHAIN: C, D;	COMPLEX (TRANSFERASE/PEPTIDE) COMPLEX (TRANSFERASE/PEPTIDE)
692	1aot	F	156	236	2e-05	-0.07	0.60		FYN PROTEIN-TYROSINE KINASE; CHAIN: F; PHOSPHOTYROSYL PEPTIDE; CHAIN: P	COMPLEX (PROTO-ONCOGENE/EARLY PROTEIN) SRC HOMOLOGY 2 DOMAIN; SH2 DOMAIN, SIGNAL TRANSDUCTION, PEPTIDE COMPLEX, 2 COMPLEX (PROTO-ONCOGENE/EARLY PROTEIN)
692	1aya	A	156	236	8e-10	0.15	0.84		HYDROLASE(SH2 DOMAIN) TYROSINE PHOSPHATASE SYP (N-TERMINAL SH2 DOMAIN) 1AYA 3 (PTP1D, SHPTP2) (E.C.3.1.3.48) COMPLEXED WITH THE PEPTIDE 1AYA 4 PDGFR-1009 1AYA 5	
692	1bkl		156	236	2e-05	0.18	0.77		PP60 V-SRC TYROSINE KINASE TRANSFORMING PROTEIN; CHAIN: NULL;	V-SRC SH2 DOMAIN SRC SH2; V-SRC SH2 DOMAIN, PHOSPHOTYROSINE RECOGNITION DOMAIN, PP60 2 SRC SH2 DOMAIN
692	1blj		156	236	1.8e-05	0.11	0.28		P55 BLK PROTEIN TYROSINE KINASE; CHAIN: NULL;	PHOSPHORYLATION SIGNAL TRANSDUCTION, TYROSINE KINASE, TRANSFERASE, 2 PHOSPHOTRANSFERASE, PHOSPHORYLATION
692	1btn		23	109	0.0016	-0.14	0.06		BETA-SPECTRIN; 1BTN 4 CHAIN: NULL; 1BTN 5	SIGNAL TRANSDUCTION PROTEIN
692	1cwd	L	156	236	0.00012	0.26	0.69		P56LCK TYROSINE KINASE; CHAIN: L; PHOSPHONOPEPTIDE CHAIN: P;	COMPLEX (PHOSPHOTRANSFERASE/PEPTIDE) PHOSPHOTRANSFERASE, COMPLEX (PHOSPHOTRANSFERASE/PEPTIDE)
692	1fao	A	18	117	4e-06	0.35	0.01		DUAL ADAPTOR OF PHOSPHOTYROSINE AND 3- CHAIN: A;	SIGNALING PROTEIN DAPP1, PHISH, BAM32; PLECKSTRIN, 3-PHOSPHOINOSITIDES, INOSITOL TETRAKISPHOSPHATE 2 SIGNAL TRANSDUCTION PROTEIN, ADAPTOR PROTEIN
692	1fb8	A	23	109	4e-06	0.51	0.64		DUAL ADAPTOR OF	SIGNALING PROTEIN DAPP1, PHISH,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									PHOSPHOTYROSINE AND 3- CHAIN: A;	BAM32; PLECKSTRIN, 3-PHOSPHOINOSITIDES, INOSITOL TETRAKISPHOSPHATE 2 SIGNAL TRANSDUCTION PROTEIN, ADAPTOR PROTEIN
692	1fhs		156	236	2e-08	0.10	0.93		GROWTH FACTOR RECEPTOR BOUND PROTEIN-2; CHAIN: NULL;	SH2 DOMAIN GRB2; GRB2, SH2 DOMAIN, PROTEIN NMR, SOLUTION STRUCTURES
692	1pls		18	117	1e-06	0.14	0.16		PHOSPHORYLATION PLECKSTRIN (N-TERMINAL PLECKSTRIN HOMOLOGY DOMAIN) MUTANT 1PLS 3 WITH LEU GLU (HIS)6 ADDED TO THE C TERMINUS 1PLS 4 (INS(G105-LEHHHHHH)) (NMR, 25 STRUCTURES) 1PLS 5	
692	1sha	A	156	236	2e-05	0.16	0.96		PHOSPHOTRANSFERASE V-SRC TYROSINE KINASE TRANSFORMING PROTEIN (PHOSPHOTYROSINE 1SHA 3 RECOGNITION DOMAIN SH2) (E.C.2.7.1.112) COMPLEX WITH 1SHA 4 PHOSPHOPEPTIDE A (TYR-VAL-PRO-MET-LEU, PHOSPHORYLATED TYR) 1SHA 5	
692	2pld	A	156	236	1.8e-06	-0.02	0.31		PHOSPHORIC DIESTER HYDROLASE PHOSPHOLIPASE C-GAMMA-1 (E.C.3.1.4.11) (C-TERMINAL SH2 2PLD 3 DOMAIN COMPRISING RESIDUES 663 - 759) COMPLEXED WITH A 2PLD 4 PHOSPHOPEPTIDE FROM THE PLATELET-DERIVED GROWTH FACTOR 2PLD 5 RECEPTOR (RESIDUES 1018 - 1029: ASP-ASN-ASP-PTYR-ILE-ILE- 2PLD 6 PRO-LEU-PRO-ASP-PRO-LYS) (NMR, MINIMIZED AVERAGE STRUCTURE) 2PLD 7	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
693	1a09	A	156	236	4e-05	0.12	0.63		C-SRC TYROSINE KINASE; CHAIN: A, B; ACE-FORMYL PHOSPHOTYR-GLU-(N,N-DIPENTYL AMINE); CHAIN: C, D;	COMPLEX (TRANSFERASE/PEPTIDE) COMPLEX (TRANSFERASE/PEPTIDE)
693	1aot	F	156	236	2e-05	-0.07	0.60		FYN PROTEIN-TYROSINE KINASE; CHAIN: F; PHOSPHOTYROSYL PEPTIDE; CHAIN: P	COMPLEX (PROTO-ONCOGENE/EARLY PROTEIN) SRC HOMOLOGY 2 DOMAIN; SH2 DOMAIN, SIGNAL TRANSDUCTION, PEPTIDE COMPLEX, 2 COMPLEX (PROTO-ONCOGENE/EARLY PROTEIN)
693	1aya	A	156	236	8e-10	0.15	0.84		HYDROLASE(SH2 DOMAIN) TYROSINE PHOSPHATASE SYP (N-TERMINAL SH2 DOMAIN) 1AYA 3 (PTP1D, SHPTP2) (E.C.3.1.3.48) COMPLEXED WITH THE PEPTIDE 1AYA 4 PDGFR-1009 1AYA 5	
693	1bkl		156	236	2e-05	0.18	0.77		PP60 V-SRC TYROSINE KINASE TRANSFORMING PROTEIN; CHAIN: NULL;	V-SRC SH2 DOMAIN SRC SH2; V-SRC SH2 DOMAIN, PHOSPHOTYROSINE RECOGNITION DOMAIN, PP60 2 SRC SH2 DOMAIN
693	1blj		156	236	1.8e-05	0.11	0.28		P55 BLK PROTEIN TYROSINE KINASE; CHAIN: NULL;	PHOSPHORYLATION SIGNAL TRANSDUCTION, TYROSINE KINASE, TRANSFERASE, 2 PHOSPHOTRANSFERASE, PHOSPHORYLATION
693	1btn		23	109	0.0016	-0.14	0.06		BETA-SPECTRIN; 1BTN 4 CHAIN: NULL; 1BTN 5	SIGNAL TRANSDUCTION PROTEIN
693	1cwd	L	156	236	0.00012	0.26	0.69		P56LCK TYROSINE KINASE; CHAIN: L; PHOSPHONOPEPTIDE CHAIN: P;	COMPLEX (PHOSPHOTRANSFERASE/PEPTIDE) PHOSPHOTRANSFERASE, COMPLEX (PHOSPHOTRANSFERASE/PEPTIDE)
693	1fao	A	18	117	4e-06	0.35	0.01		DUAL ADAPTOR OF PHOSPHOTYROSINE AND 3-CHAIN: A;	SIGNALING PROTEIN DAPP1, PHISH, BAM32; PLECKSTRIN, 3-PHOSPHOINOSITIDES, INOSITOL TETRAKISPHOSPHATE 2 SIGNAL TRANSDUCTION PROTEIN, ADAPTOR PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
693	1fb8	A	23	109	4e-06	0.51	0.64		DUAL ADAPTOR OF PHOSPHOTYROSINE AND 3-CHAIN: A;	SIGNALING PROTEIN DAPP1, PHISH, BAM32; PLECKSTRIN, 3-PHOSPHOINOSITIDES, INOSITOL TETRAKISPHOSPHATE 2 SIGNAL TRANSDUCTION PROTEIN, ADAPTOR PROTEIN
693	1fhs		156	236	2e-08	0.10	0.93		GROWTH FACTOR RECEPTOR BOUND PROTEIN-2; CHAIN: NULL;	SH2 DOMAIN GRB2; GRB2, SH2 DOMAIN, PROTEIN NMR, SOLUTION STRUCTURES
693	1pls		18	117	1e-06	0.14	0.16		PHOSPHORYLATION PLECKSTRIN (N-TERMINAL PLECKSTRIN HOMOLOGY DOMAIN) MUTANT 1PLS 3 WITH LEU GLU (HIS)6 ADDED TO THE C TERMINUS 1PLS 4 (INS(G105-LEHHHHHH)) (NMR, 25 STRUCTURES) 1PLS 5	
693	1sha	A	156	236	2e-05	0.16	0.96		PHOSPHOTRANSFERASE V-SRC TYROSINE KINASE TRANSFORMING PROTEIN (PHOSPHOTYROSINE 1SHA 3 RECOGNITION DOMAIN SH2) (E.C.2.7.1.112) COMPLEX WITH 1SHA 4 PHOSPHOPEPTIDE A (TYR-VAL-PRO-MET-LEU, PHOSPHORYLATED TYR) 1SHA 5	
693	2pld	A	156	236	1.8e-06	-0.02	0.31		PHOSPHORIC DIESTER HYDROLASE PHOSPHOLIPASE C-GAMMA-1 (E.C.3.1.4.11) (C-TERMINAL SH2 2PLD 3 DOMAIN COMPRISING RESIDUES 663 - 759) COMPLEXED WITH A 2PLD 4 PHOSPHOPEPTIDE FROM THE PLATELET-DERIVED GROWTH FACTOR 2PLD 5 RECEPTOR (RESIDUES 1018 - 1029: ASP-ASN-ASP-PTYR-ILE-ILE- 2PLD 6 PRO-LEU-PRO-ASP-PRO-LYS) (NMR, MINIMIZED AVERAGE	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									STRUCTURE) 2PLD 7	
693	1a09	A	156	236	4e-05	0.12	0.63		C-SRC TYROSINE KINASE; CHAIN: A, B; ACE-FORMYL PHOSPHOTYR-GLU-(N,N-DIPENTYL AMINE); CHAIN: C, D;	COMPLEX (TRANSFERASE/PEPTIDE) COMPLEX (TRANSFERASE/PEPTIDE)
693	1aot	F	156	236	2e-05	-0.07	0.60		FYN PROTEIN-TYROSINE KINASE; CHAIN: F; PHOSPHOTYROSYL PEPTIDE; CHAIN: P	COMPLEX (PROTO-ONCOGENE/EARLY PROTEIN) SRC HOMOLOGY 2 DOMAIN; SH2 DOMAIN, SIGNAL TRANSDUCTION, PEPTIDE COMPLEX, 2 COMPLEX (PROTO-ONCOGENE/EARLY PROTEIN)
693	1aya	A	156	236	8e-10	0.15	0.84		HYDROLASE(SH2 DOMAIN) TYROSINE PHOSPHATASE SYP (N-TERMINAL SH2 DOMAIN) 1AYA 3 (PTP1D, SHPTP2) (E.C.3.1.3.48) COMPLEXED WITH THE PEPTIDE 1AYA 4 PDGFR-1009 1AYA 5	
693	1bkl		156	236	2e-05	0.18	0.77		PP60 V-SRC TYROSINE KINASE TRANSFORMING PROTEIN; CHAIN: NULL;	V-SRC SH2 DOMAIN SRC SH2; V-SRC SH2 DOMAIN, PHOSPHOTYROSINE RECOGNITION DOMAIN, PP60 2 SRC SH2 DOMAIN
693	1blj		156	236	1.8e-05	0.11	0.28		P55 BLK PROTEIN TYROSINE KINASE; CHAIN: NULL;	PHOSPHORYLATION SIGNAL TRANSDUCTION, TYROSINE KINASE, TRANSFERASE, 2 PHOSPHOTRANSFERASE, PHOSPHORYLATION
693	1btn		23	109	0.0016	-0.14	0.06		BETA-SPECTRIN; 1BTN 4 CHAIN: NULL; 1BTN 5	SIGNAL TRANSDUCTION PROTEIN
693	1cwd	L	156	236	0.00012	0.26	0.69		P56LCK TYROSINE KINASE; CHAIN: L; PHOSPHONOPEPTIDE CHAIN: P;	COMPLEX (PHOSPHOTRANSFERASE/PEPTIDE) PHOSPHOTRANSFERASE, COMPLEX (PHOSPHOTRANSFERASE/PEPTIDE)
693	1fao	A	18	117	4e-06	0.35	0.01		DUAL ADAPTOR OF PHOSPHOTYROSINE AND 3-CHAIN: A;	SIGNALING PROTEIN DAPP1, PHISH, BAM32; PLECKSTRIN, 3-PHOSPHOINOSITIDES, INOSITOL TETRAKISPHOSPHATE 2 SIGNAL TRANSDUCTION PROTEIN, ADAPTOR PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
693	1fb8	A	23	109	4e-06	0.51	0.64		DUAL ADAPTOR OF PHOSPHOTYROSINE AND 3-CHAIN: A;	SIGNALING PROTEIN DAPP1, PHISH, BAM32; PLECKSTRIN, 3-PHOSPHOINOSITIDES, INOSITOL TETRAKISPHOSPHATE 2 SIGNAL TRANSDUCTION PROTEIN, ADAPTOR PROTEIN
693	1fhs		156	236	2e-08	0.10	0.93		GROWTH FACTOR RECEPTOR BOUND PROTEIN-2; CHAIN: NULL;	SH2 DOMAIN GRB2; GRB2, SH2 DOMAIN, PROTEIN NMR, SOLUTION STRUCTURES
693	1pls		18	117	1e-06	0.14	0.16		PHOSPHORYLATION PLECKSTRIN (N-TERMINAL PLECKSTRIN HOMOLOGY DOMAIN) MUTANT 1PLS 3 WITH LEU GLU (HIS)6 ADDED TO THE C TERMINUS 1PLS 4 (INS(G105-LEHHHHHH)) (NMR, 25 STRUCTURES) 1PLS 5	
693	1sha	A	156	236	2e-05	0.16	0.96		PHOSPHOTRANSFERASE V-SRC TYROSINE KINASE TRANSFORMING PROTEIN (PHOSPHOTYROSINE 1SHA 3 RECOGNITION DOMAIN SH2) (E.C.2.7.1.112) COMPLEX WITH 1SHA 4 PHOSPHOPEPTIDE A (TYR-VAL-PRO-MET-LEU, PHOSPHORYLATED TYR) 1SHA 5	
693	2pld	A	156	236	1.8e-06	-0.02	0.31		PHOSPHORIC DIESTER HYDROLASE PHOSPHOLIPASE C-GAMMA-1 (E.C.3.1.4.11) (C-TERMINAL SH2 2PLD 3 DOMAIN COMPRISING RESIDUES 663 - 759) COMPLEXED WITH A 2PLD 4 PHOSPHOPEPTIDE FROM THE PLATELET-DERIVED GROWTH FACTOR 2PLD 5 RECEPTOR (RESIDUES 1018 - 1029: ASP-ASN-ASP-PTYR-ILE-ILE- 2PLD 6 PRO-LEU-PRO-ASP-PRO-LYS) (NMR, MINIMIZED AVERAGE	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									STRUCTURE) 2PLD 7	
694	1sfp		94	208	1.6e-17	0.59	0.89		ASFP; CHAIN: NULL;	SPERMADHESIN ACIDIC SEMINAL PROTEIN; SPERMADHESIN, BOVINE SEMINAL PLASMA PROTEIN, ACIDIC 2 SEMINAL FLUID PROTEIN, ASFP, CUB DOMAIN, X-RAY CRYSTAL 3 STRUCTURE, GROWTH FACTOR
694	1sfp		99	206	2e-19	0.57	0.68		ASFP; CHAIN: NULL;	SPERMADHESIN ACIDIC SEMINAL PROTEIN; SPERMADHESIN, BOVINE SEMINAL PLASMA PROTEIN, ACIDIC 2 SEMINAL FLUID PROTEIN, ASFP, CUB DOMAIN, X-RAY CRYSTAL 3 STRUCTURE, GROWTH FACTOR
694	1spp	A	97	208	1.1e-14	0.15	0.35		MAJOR SEMINAL PLASMA GLYCOPROTEIN PSP-I; CHAIN: A; MAJOR SEMINAL PLASMA GLYCOPROTEIN PSP-II; CHAIN: B	COMPLEX (SEMINAL PLASMA PROTEIN/SPP) SEMINAL PLASMA PROTEINS, SPERMADHESINS, CUB DOMAIN 2 ARCHITECTURE, COMPLEX (SEMINAL PLASMA PROTEIN/SPP)
694	1spp	A	98	206	4e-20	0.26	0.43		MAJOR SEMINAL PLASMA GLYCOPROTEIN PSP-I; CHAIN: A; MAJOR SEMINAL PLASMA GLYCOPROTEIN PSP-II; CHAIN: B	COMPLEX (SEMINAL PLASMA PROTEIN/SPP) SEMINAL PLASMA PROTEINS, SPERMADHESINS, CUB DOMAIN 2 ARCHITECTURE, COMPLEX (SEMINAL PLASMA PROTEIN/SPP)
694	1spp	B	96	204	3.6e-15	0.56	0.45		MAJOR SEMINAL PLASMA GLYCOPROTEIN PSP-I; CHAIN: A; MAJOR SEMINAL PLASMA GLYCOPROTEIN PSP-II; CHAIN: B	COMPLEX (SEMINAL PLASMA PROTEIN/SPP) SEMINAL PLASMA PROTEINS, SPERMADHESINS, CUB DOMAIN 2 ARCHITECTURE, COMPLEX (SEMINAL PLASMA PROTEIN/SPP)
694	1spp	B	99	203	2e-21	0.43	0.25		MAJOR SEMINAL PLASMA GLYCOPROTEIN PSP-I; CHAIN: A; MAJOR SEMINAL PLASMA GLYCOPROTEIN PSP-II; CHAIN: B	COMPLEX (SEMINAL PLASMA PROTEIN/SPP) SEMINAL PLASMA PROTEINS, SPERMADHESINS, CUB DOMAIN 2 ARCHITECTURE, COMPLEX (SEMINAL PLASMA PROTEIN/SPP)
698	1d2h	A	87	243	1.6e-06	-0.06	0.35		GLYCINE N-METHYLTRANSFERASE; CHAIN: A, B, C, D;	TRANSFERASE METHYLTRANSFERASE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
698	1d2h	A	92	211	3.6e-16	-0.15	0.37		GLYCINE N-METHYLTRANSFERASE; CHAIN: A, B, C, D;	TRANSFERASE METHYLTRANSFERASE
698	1vid		49	262	1.8e-32			81.95	CATECHOL O-METHYLTRANSFERASE; CHAIN: NULL;	TRANSFERASE (METHYLTRANSFERASE) COMT; TRANSFERASE, METHYLTRANSFERASE, NEUROTRANSMITTER DEGRADATION
698	1vid		50	249	1.8e-32	0.43	0.81		CATECHOL O-METHYLTRANSFERASE; CHAIN: NULL;	TRANSFERASE (METHYLTRANSFERASE) COMT; TRANSFERASE, METHYLTRANSFERASE, NEUROTRANSMITTER DEGRADATION
698	1vid		52	252	9e-19	0.38	0.40		CATECHOL O-METHYLTRANSFERASE; CHAIN: NULL;	TRANSFERASE (METHYLTRANSFERASE) COMT; TRANSFERASE, METHYLTRANSFERASE, NEUROTRANSMITTER DEGRADATION
698	1xva	A	81	246	1.4e-07	-0.04	0.70		GLYCINE N-METHYLTRANSFERASE; CHAIN: A, B;	METHYLTRANSFERASE GNMT, S-ADENOSYL-L-METHIONINE\: GLYCINE METHYLTRANSFERASE
698	1xva	A	92	211	3.6e-16	-0.25	0.35		GLYCINE N-METHYLTRANSFERASE; CHAIN: A, B;	METHYLTRANSFERASE GNMT, S-ADENOSYL-L-METHIONINE\: GLYCINE METHYLTRANSFERASE
700	1vid		36	219	4e-06	0.35	0.57		CATECHOL O-METHYLTRANSFERASE; CHAIN: NULL;	TRANSFERASE (METHYLTRANSFERASE) COMT; TRANSFERASE, METHYLTRANSFERASE, NEUROTRANSMITTER DEGRADATION
701	1bor		79	119	7.2e-05	-0.69	0.05		TRANSCRIPTION FACTOR PML; CHAIN: NULL;	TRANSCRIPTION REGULATION PROTO-ONCOGENE, NUCLEAR BODIES (PODS), LEUKEMIA, 2 TRANSCRIPTION REGULATION
701	1chc		70	128	6e-18	0.02	0.30		VIRUS EQUINE HERPES VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	
701	1chc		74	132	3.6e-15	-0.18	0.27		VIRUS EQUINE HERPES VIRUS-1 (C3HC4, OR RING DOMAIN) 1CHC 3 (NMR, 1 STRUCTURE) 1CHC 4	
701	1rmd		73	124	4e-14	-0.07	0.93		RAG1; CHAIN: NULL;	DNA-BINDING PROTEIN V(D)J

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										RECOMBINATION ACTIVATING PROTEIN 1; RAG1, V(D)J RECOMBINATION, ANTIBODY, MAD, RING FINGER, 2 ZINC BINUCLEAR CLUSTER, ZINC FINGER, DNA-BINDING PROTEIN
701	1rmd		76	132	7.2e-10	0.04	0.63		RAG1; CHAIN: NULL;	DNA-BINDING PROTEIN V(D)J RECOMBINATION ACTIVATING PROTEIN 1; RAG1, V(D)J RECOMBINATION, ANTIBODY, MAD, RING FINGER, 2 ZINC BINUCLEAR CLUSTER, ZINC FINGER, DNA-BINDING PROTEIN
702	1a4y	A	145	595	2e-38	0.11	1.00		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPITOPE MAPPING, LEUCINE-RICH 3 REPEATS
702	1a4y	A	200	665	6e-39	0.09	0.57		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPITOPE MAPPING, LEUCINE-RICH 3 REPEATS
702	1a4y	A	236	610	1.4e-21	0.04	0.11		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPITOPE MAPPING, LEUCINE-RICH 3 REPEATS
702	1a4y	A	309	728	2e-42	0.31	0.94		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPITOPE MAPPING, LEUCINE-RICH 3 REPEATS
702	1a4y	A	336	689	1.8e-23	0.21	0.22		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPITOPE MAPPING, LEUCINE-RICH 3 REPEATS
702	1a4y	A	414	762	9e-22	0.11	0.12		RIBONUCLEASE INHIBITOR;	COMPLEX (INHIBITOR/NUCLEASE)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPITOPE MAPPING, LEUCINE-RICH 3 REPEATS
702	1a4y	A	495	869	5.4e-20	-0.02	0.10		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPITOPE MAPPING, LEUCINE-RICH 3 REPEATS
702	1a4y	A	64	535	2e-39	-0.05	0.27		RIBONUCLEASE INHIBITOR; CHAIN: A, D; ANGIOGENIN; CHAIN: B, E;	COMPLEX (INHIBITOR/NUCLEASE) COMPLEX (INHIBITOR/NUCLEASE), COMPLEX (RI-ANG), HYDROLASE 2 MOLECULAR RECOGNITION, EPITOPE MAPPING, LEUCINE-RICH 3 REPEATS
702	1a9n	A	119	250	2e-16	0.04	0.42		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	A	129	276	2e-16	0.30	0.58		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	A	206	374	2e-17	0.18	0.39		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	A	227	428	8e-17	0.20	0.12		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	A	315	501	1.4e-13	0.20	-0.08		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	A	397	551	1.8e-17	0.32	0.78		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	A	43	204	2e-12	-0.19	0.00		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	A	469	608	1.2e-20	0.03	0.05		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
702	1a9n	A	528	683	1.6e-16	0.08	0.37		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	A	574	707	1.8e-17	0.45	0.86		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	A	659	774	6e-13	0.30	0.04		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	C	119	250	4e-17	-0.19	0.22		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	C	206	374	2e-17	0.13	-0.07		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	C	227	428	4e-17	-0.02	0.31		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	C	315	501	6e-14	0.09	-0.09		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	C	469	608	6e-21	0.12	0.19		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	C	528	690	1e-16	0.24	0.43		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	C	574	707	4e-17	0.47	0.92		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1a9n	C	624	756	2e-17	0.24	-0.05		U2 RNA HAIRPIN IV; CHAIN: Q, R; U2 A'; CHAIN: A, C; U2 B"; CHAIN: B, D;	COMPLEX (NUCLEAR PROTEIN/RNA) COMPLEX (NUCLEAR PROTEIN/RNA), RNA, SNRNP,RIBONUCLEOPROTEIN
702	1d0b	A	12	174	7.2e-23	-0.04	0.18		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
702	1d0b	A	122	341	5.4e-20	-0.04	0.23		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
702	1d0b	A	126	373	6e-25	-0.05	0.03		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										CALCIUM BINDING, CELL ADHESION
702	1d0b	A	203	425	2e-23	-0.01	0.15		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
702	1d0b	A	304	479	7.2e-24	0.29	1.00		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
702	1d0b	A	304	532	8e-23	0.14	0.93		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
702	1d0b	A	363	552	1.6e-24	0.42	0.60		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
702	1d0b	A	401	582	5.4e-25	0.06	-0.13		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
702	1d0b	A	44	254	4e-21	-0.22	0.05		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
702	1d0b	A	476	711	6e-31	-0.00	0.16		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
702	1d0b	A	482	660	1.4e-23	0.09	0.82		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
702	1d0b	A	520	708	7.2e-24	0.38	1.00		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
702	1d0b	A	652	801	1.8e-24	0.27	0.65		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
702	1d0b	A	687	862	5.4e-20	0.04	-0.17		INTERNALIN B; CHAIN: A;	CELL ADHESION LEUCINE RICH REPEAT, CALCIUM BINDING, CELL ADHESION
702	1dce	A	193	526	1.2e-12	-0.17	0.01		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
702	1dce	A	217	439	2e-16	0.17	0.25		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
702	1dce	A	247	362	5.4e-12	0.30	0.10		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A,	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
702	1dce	A	287	403	5.4e-10	0.40	0.81		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
702	1dce	A	519	737	1.6e-14	-0.03	0.34		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
702	1dce	A	548	665	3.6e-13	-0.24	0.53		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
702	1dce	A	656	762	1.8e-11	0.38	0.43		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
702	1ds9	A	124	314	1.8e-19	-0.16	0.11		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
702	1ds9	A	296	447	5.4e-14	0.12	0.27		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
702	1ds9	A	46	155	2e-11	-0.31	0.62		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										FLAGELLA
702	1ds9	A	477	634	1.6e-18	-0.34	0.47		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
702	1ds9	A	484	634	5.4e-15	-0.29	0.06		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
702	1ds9	A	513	683	1.4e-17	-0.15	0.07		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
702	1ds9	A	533	683	3.6e-12	-0.26	0.04		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
702	1ds9	A	627	784	6e-16	-0.19	0.15		OUTER ARM DYNEIN; CHAIN: A;	CONTRACTILE PROTEIN LEUCINE-RICH REPEAT, BETA-BETA-ALPHA CYLINDER, DYNEIN, 2 CHLAMYDOMONAS, FLAGELLA
702	1fo1	A	306	376	1.8e-06	-0.01	0.47		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
702	1fo1	B	306	376	1.8e-06	-0.57	0.10		NUCLEAR RNA EXPORT FACTOR 1; CHAIN: A, B;	RNA BINDING PROTEIN TAP (NFX1); RIBONUCLEOPROTEIN (RNP,RBD OR RRM) AND LEUCINE-RICH-REPEAT 2 (LRR)
702	1fqv	A	263	497	3.6e-14	0.18	-0.15		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
702	1fqv	A	459	732	1.6e-23	-0.01	0.03		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
702	1fqv	A	578	801	1.8e-09	-0.00	-0.07		SKP2; CHAIN: A, C, E, G, I, K, M, O; SKP1; CHAIN: B, D, F, H, J, L, N, P;	LIGASE CYCLIN A/CDK2-ASSOCIATED PROTEIN P45; CYCLIN A/CDK2-ASSOCIATED PROTEIN P19; SKP1, SKP2, F-BOX, LRR, LEUCINE-RICH REPEAT, SCF, UBIQUITIN, 2 E3, UBIQUITIN PROTEIN LIGASE
702	1fs2	A	240	420	5.4e-09	0.09	0.22		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
702	1fs2	A	263	475	3.6e-13	0.19	-0.14		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
702	1fs2	A	317	606	4e-20	0.10	0.29		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
702	1fs2	A	367	550	1.1e-13	0.06	0.06		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
702	1fs2	A	479	709	2e-30	0.07	-0.07		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
702	1fs2	A	499	705	1.8e-12	0.28	0.13		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										UBIQUITIN PROTEIN LIGASE
702	1fs2	A	546	750	5.4e-11	0.24	0.16		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
702	1fs2	A	567	758	6e-19	0.08	0.12		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
702	1fs2	A	64	342	4e-20	-0.29	0.07		SKP2; CHAIN: A, C; SKP1; CHAIN: B, D;	LIGASE CYCLIN A/CDK2-ASSOCIATED P45; CYCLIN A/CDK2-ASSOCIATED P19; SKP1, SKP2, F-BOX, LRRS, LEUCINE-RICH REPEATS, SCF, 2 UBIQUITIN, E3, UBIQUITIN PROTEIN LIGASE
702	1yrg	A	239	428	9e-11	-0.11	0.01		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI1, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE- 2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
702	1yrg	A	279	485	1.8e-13	0.30	0.01		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI1, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE- 2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
702	1yrg	A	477	738	2e-30	0.10	0.13		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI1, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE- 2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
702	1yrg	A	547	774	2e-19	0.43	0.16		GTPASE-ACTIVATING PROTEIN	TRANSCRIPTION RNA1P; RANGAP;

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									RNA1_SCHPO; CHAIN: A, B;	GTPASE-ACTIVATING PROTEIN FOR SPI1, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE- 2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
702	1yrg	A	64	254	6e-19	0.11	0.31		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI1, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE- 2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
702	1yrg	A	64	344	4e-18	0.08	0.12		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI1, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE- 2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
702	1yrg	A	88	428	4e-21	0.09	0.19		GTPASE-ACTIVATING PROTEIN RNA1_SCHPO; CHAIN: A, B;	TRANSCRIPTION RNA1P; RANGAP; GTPASE-ACTIVATING PROTEIN FOR SPI1, GTPASE-ACTIVATING PROTEIN, GAP, RNA1P, RANGAP, LRR, LEUCINE- 2 RICH REPEAT PROTEIN, TWINNING, HEMIHEDRAL TWINNING, 3 MEROHEDRAL TWINNING, MEROHEDRY
702	2bnh		204	659	6e-53	0.03	0.75		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
702	2bnh		313	757	8e-49	0.27	0.87		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
702	2bnh		315	728	3.6e-26	0.01	0.53		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
702	2bnh		477	869	3.6e-24	0.02	0.53		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										ACETYLATION, LEUCINE-RICH REPEATS
702	2bnh		67	547	1.2e-45	-0.10	0.53		RIBONUCLEASE INHIBITOR; CHAIN: NULL;	ACETYLATION RNASE INHIBITOR, RIBONUCLEASE/ANGIOGENIN INHIBITOR ACETYLATION, LEUCINE-RICH REPEATS
703	1ps2		40	94	1.8e-17			61.22	PS2; CHAIN: NULL;	GROWTH FACTOR PNR-2; GROWTH FACTOR, CELL MOTILITY, TUMOR SUPPRESSOR, TREFOIL 2 DOMAIN, SIGNAL
703	1ps2		43	86	1.8e-17	0.38	1.00		PS2; CHAIN: NULL;	GROWTH FACTOR PNR-2; GROWTH FACTOR, CELL MOTILITY, TUMOR SUPPRESSOR, TREFOIL 2 DOMAIN, SIGNAL
703	2psp	A	2	94	1.8e-19			60.69	PORCINE PANCREATIC SPASMOLYTIC POLYPEPTIDE; CHAIN: A, B;	TREFOIL FAMILY OF PEPTIDES PSP REPEAT, GROWTH FACTOR, SIGNAL
703	2psp	A	43	94	1.8e-19	0.17	1.00		PORCINE PANCREATIC SPASMOLYTIC POLYPEPTIDE; CHAIN: A, B;	TREFOIL FAMILY OF PEPTIDES PSP REPEAT, GROWTH FACTOR, SIGNAL
706	1c4x	A	181	296	0.00018	0.03	0.01		2-HYDROXY-6-OXO-6-PHENYLHEXA-2,4-DIENOATE CHAIN: A;	HYDROLASE BPHD; HYDROLASE, PCB DEGRADATION
706	1c7j	A	43	609	1.3e-97	0.28	0.87		PARA-NITROBENZYL ESTERASE; CHAIN: A;	HYDROLASE PNB ESTERASE; ALPHA-BETA HYDROLASE, DIRECTED EVOLUTION, ORGANIC ACTIVITY, 2 PNB ESTERASE
706	1cle	A	41	584	3.6e-79			161.50	CHOLESTEROL ESTERASE; 1CLE 4 CHAIN: A, B; 1CLE 5	LIPASE ESTERASE, SUBSTRATE/PRODUCT-BOUND 1CLE 9
706	1cle	A	67	563	3.6e-79	0.23	0.96		CHOLESTEROL ESTERASE; 1CLE 4 CHAIN: A, B; 1CLE 5	LIPASE ESTERASE, SUBSTRATE/PRODUCT-BOUND 1CLE 9
706	1din		179	296	0.009	0.15	0.24		DIENELACTONE HYDROLASE; CHAIN: NULL;	HYDROLYTIC ENZYME DLH; DIENELACTONE HYDROLASE, AROMATIC HYDROCARBON CATABOLISM, 2 SERINE ESTERASE, CARBOXYMETHYLENEBUTENOLIDASE, 3 HYDROLYTIC ENZYME
706	1dx4	A	40	614	0	0.41	1.00		ACETYLCHOLINESTERASE;	HYDROLASE (SERINE ESTERASE)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: A;	HYDROLASE (SERINE ESTERASE), HYDROLASE, SERINE ESTERASE, 2 SYNAPSE, MEMBRANE, NERVE, MUSCLE, SIGNAL, NEUROTRANSMITTER 3 DEGRADATION, GLYCOPROTEIN, GPI-ANCHOR, ALTERNATIVE SPLICING
706	1ea5	A	40	615	0	0.22	1.00		ACETYLCHOLINESTERASE; CHAIN: A;	CHOLINESTERASE SERINE HYDROLASE, NEUROTRANSMITTER CLEAVAGE, CATALYTIC 2 TRIAD, ALPHA/BETA HYDROLASE
706	1evq	A	179	349	5.4e-28	0.10	0.43		SERINE HYDROLASE; CHAIN: A;	HYDROLASE ALPHA/BETA HYDROLASE FOLD
706	1evq	A	73	372	2e-40	-0.08	0.24		SERINE HYDROLASE; CHAIN: A;	HYDROLASE ALPHA/BETA HYDROLASE FOLD
706	1f6w	A	44	615	0	0.38	1.00		BILE SALT ACTIVATED LIPASE; CHAIN: A;	HYDROLASE BILE SALT ACTIVATED LIPASE, ESTERASE, CATALYTIC DOMAIN
706	1jkm	A	180	337	5.4e-16	0.24	0.49		BREFELDIN A ESTERASE; CHAIN: A, B;	SERINE HYDROLASE SERINE HYDROLASE, DEGRADATION OF BREFELDIN A, ALPHA/BETA 2 HYDROLASE FAMILY
706	1lpp		41	584	3.6e-78			167.96	HYDROLASE LIPASE (E.C.3.1.1.3) (TRIACYLGLYCEROL LIPASE) COMPLEXED WITH 1LPP 3 HEXADECANESULFONATE 1LPP 4 1LPP 71	
706	1lpp		67	563	3.6e-78	0.32	0.93		HYDROLASE LIPASE (E.C.3.1.1.3) (TRIACYLGLYCEROL LIPASE) COMPLEXED WITH 1LPP 3 HEXADECANESULFONATE 1LPP 4 1LPP 71	
706	1maa	A	38	615	0	0.48	1.00		ACETYLCHOLINESTERASE; CHAIN: A, B, C, D;	HYDROLASE MACHE; HYDROLASE, SERINE ESTERASE, ACETYLCHOLINESTERASE, TETRAMER, 2 HYDROLASE FOLD, GLYCOSYLATED PROTEIN
706	1maa	A	38	615	0			364.47	ACETYLCHOLINESTERASE; CHAIN: A, B, C, D;	HYDROLASE MACHE; HYDROLASE, SERINE ESTERASE, ACETYLCHOLINESTERASE, TETRAMER, 2

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										HYDROLASE FOLD, GLYCOSYLATED PROTEIN
706	1qe3	A	40	600	3.6e-93			223.02	PARA-NITROBENZYL ESTERASE; CHAIN: A;	HYDROLASE PNB ESTERASE; ALPHA-BETA HYDROLASE DIRECTED EVOLUTION
706	1qe3	A	43	602	3.6e-93	0.28	0.99		PARA-NITROBENZYL ESTERASE; CHAIN: A;	HYDROLASE PNB ESTERASE; ALPHA-BETA HYDROLASE DIRECTED EVOLUTION
706	1qfm	A	181	342	1.3e-21	0.31	0.31		PROLYL OLIGOPEPTIDASE; CHAIN: A;	HYDROLASE PROLYL ENDOPEPTIDASE, POST-PROLINE CLEAVING PROLYL OLIGOPEPTIDASE, AMNESIA, ALPHA/BETA-HYDROLASE, BETA-2 PROPELLER
706	1qfm	A	35	393	2e-54	0.14	0.30		PROLYL OLIGOPEPTIDASE; CHAIN: A;	HYDROLASE PROLYL ENDOPEPTIDASE, POST-PROLINE CLEAVING PROLYL OLIGOPEPTIDASE, AMNESIA, ALPHA/BETA-HYDROLASE, BETA-2 PROPELLER
706	1qtr	A	184	284	9e-05	-0.18	0.03		PROLYL AMINOPEPTIDASE; CHAIN: A;	HYDROLASE ALPHA BETA HYDROLASE FOLD, PROLINE, PROLYL AMINOPEPTIDASE, 2 SERRATIA, IMINOPEPTIDASE
706	1thg		45	583	5.4e-86			196.76	HYDROLASE(CARBOXYLIC ESTERASE) LIPASE (E.C.3.1.1.3) TRIACYLGLYCEROL HYDROLASE 1THG 3	
706	1thg		47	566	5.4e-86	0.40	1.00		HYDROLASE(CARBOXYLIC ESTERASE) LIPASE (E.C.3.1.1.3) TRIACYLGLYCEROL HYDROLASE 1THG 3	
706	2bce		39	619	0			302.27	CHOLESTEROL ESTERASE; CHAIN: NULL;	HYDROLASE BILE SALT ACTIVATED LIPASE, BILE SALT STIMULATED HYDROLASE, SERINE ESTERASE, LIPASE
706	2bce		44	615	0	0.43	1.00		CHOLESTEROL ESTERASE; CHAIN: NULL;	HYDROLASE BILE SALT ACTIVATED LIPASE, BILE SALT STIMULATED HYDROLASE, SERINE ESTERASE, LIPASE
710	1a06		12	322	1.4e-84			125.52	CALCIUM/CALMODULIN-	KINASE KINASE, SIGNAL

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									DEPENDENT PROTEIN KINASE; CHAIN: NULL;	TRANSDUCTION, CALCIUM/CALMODULIN
710	1a06		23	311	1.4e-84	0.03	1.00		CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE; CHAIN: NULL;	KINASE KINASE, SIGNAL TRANSDUCTION, CALCIUM/CALMODULIN
710	1a6o		1	335	7.2e-43			96.62	PROTEIN KINASE CK2/ALPHA-SUBUNIT; CHAIN: NULL;	TRANSFERASE TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, CASEIN KINASE, 2 SER/THR KINASE
710	1apm	E	13	317	0	0.56	1.00		TRANSFERASE(PHOSPHOTRANS FERASE) $C-/AMP$-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (C/APK) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (/S139A$) COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	
710	1apm	E	1	334	0			214.61	TRANSFERASE(PHOSPHOTRANS FERASE) $C-/AMP$-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (C/APK) 1APM 3 (CATALYTIC SUBUNIT) ALPHA ISOENZYME MUTANT WITH SER 139 1APM 4 REPLACED BY ALA (/S139A$) COMPLEX WITH THE PEPTIDE 1APM 5 INHIBITOR PKI(5-24) AND THE DETERGENT MEGA-8 1APM 6	
710	1aq1		20	286	3.6e-55	0.14	1.00		CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE, PHOSPHORYLATION, STAUROSPORINE, 2 CELL DIVISION, MITOSIS, INHIBITION
710	1aq1		20	329	3.6e-55			102.35	CYCLIN-DEPENDENT PROTEIN KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; PROTEIN KINASE, CELL CYCLE, PHOSPHORYLATION, STAUROSPORINE, 2 CELL DIVISION, MITOSIS, INHIBITION

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
710	1bi8	A	21	314	1.1e-43			87.13	CYCLIN-DEPENDENT KINASE 6; CHAIN: A, C; CYCLIN-DEPENDENT KINASE INHIBITOR; CHAIN: B, D;	COMPLEX (KINASE/INHIBITOR) CDK6; P19INK4D; CYCLIN DEPENDENT KINASE, CYCLIN DEPENDENT KINASE INHIBITORY 2 PROTEIN, CDK, INK4, CELL CYCLE, COMPLEX (KINASE/INHIBITOR) HEADER HELIX
710	1blx	A	15	335	3.6e-47			103.01	CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
710	1blx	A	23	285	3.6e-47	0.41	1.00		CYCLIN-DEPENDENT KINASE 6; CHAIN: A; P19INK4D; CHAIN: B;	COMPLEX (INHIBITOR PROTEIN/KINASE) INHIBITOR PROTEIN, CYCLIN-DEPENDENT KINASE, CELL CYCLE 2 CONTROL, ALPHA/BETA, COMPLEX (INHIBITOR PROTEIN/KINASE)
710	1byg	A	15	288	7.2e-31			85.43	C-TERMINAL SRC KINASE; CHAIN: A;	TRANSFERASE CSK; PROTEIN KINASE, C-TERMINAL SRC KINASE, PHOSPHORYLATION, 2 STAUROSPORINE, TRANSFERASE
710	1cki	A	16	312	4e-51			70.32	CASEIN KINASE I DELTA; 1CKI 6 CHAIN: A, B; 1CKI 7	PHOSPHOTRANSFERASE PROTEIN KINASE 1CKI 18
710	1cki	A	21	292	4e-51	-0.11	0.84		CASEIN KINASE I DELTA; 1CKI 6 CHAIN: A, B; 1CKI 7	PHOSPHOTRANSFERASE PROTEIN KINASE 1CKI 18
710	1cm8	A	37	302	7.2e-45	0.31	1.00		PHOSPHORYLATED MAP KINASE P38-GAMMA; CHAIN: A, B;	TRANSFERASE STRESS-ACTIVATED PROTEIN KINASE-3, ERK6, ERK5; P38-GAMMA, GAMMA, PHOSPHORYLATION, MAP KINASE
710	1cmk	E	13	317	0	0.68	1.00		PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	
710	1cmk	E	2	334	0			213.80	PHOSPHOTRANSFERASE CAMP-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT 1CMK 3 (E.C.2.7.1.37) 1CMK 4	
710	1csn		17	318	4e-52			75.78	CASEIN KINASE-1; 1CSN 4	PHOSPHOTRANSFERASE
710	1csn		22	293	4e-52	0.33	0.92		CASEIN KINASE-1; 1CSN 4	PHOSPHOTRANSFERASE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
710	1ctp	E	1	325	0			212.30	TRANSFERASE(PHOSPHOTRANSFERASE) CAMP-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (CAPK) 1CTP 3 (CATALYTIC SUBUNIT) 1CTP 4	
710	1ctp	E	13	317	0	0.58	1.00		TRANSFERASE(PHOSPHOTRANSFERASE) CAMP-DEPENDENT PROTEIN KINASE (E.C.2.7.1.37) (CAPK) 1CTP 3 (CATALYTIC SUBUNIT) 1CTP 4	
710	1f3m	C	21	298	1.8e-66	0.22	1.00		SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: A, B; SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: C, D;	TRANSFERASE KINASE DOMAIN, AUTOINHIBITORY FRAGMENT, HOMODIMER
710	1f3m	C	22	284	1.8e-58	0.22	1.00		SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: A, B; SERINE/THREONINE-PROTEIN KINASE PAK-ALPHA; CHAIN: C, D;	TRANSFERASE KINASE DOMAIN, AUTOINHIBITORY FRAGMENT, HOMODIMER
710	1fgk	A	14	287	4e-32			99.10	FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
710	1fgk	B	5	287	1.1e-37			108.19	FGF RECEPTOR 1; CHAIN: A, B;	PHOSPHOTRANSFERASE FGFR1K, FIBROBLAST GROWTH FACTOR RECEPTOR 1; TRANSFERASE, TYROSINE-PROTEIN KINASE, ATP-BINDING, 2 PHOSPHORYLATION, RECEPTOR, PHOSPHOTRANSFERASE
710	1hcl		20	286	1.3e-57	0.34	1.00		HUMAN CYCLIN-DEPENDENT KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; TRANSFERASE, SERINE/THREONINE PROTEIN KINASE, ATP-BINDING, 2 CELL CYCLE, CELL DIVISION, MITOSIS, PHOSPHORYLATION
710	1hcl		20	329	1.3e-57			112.61	HUMAN CYCLIN-DEPENDENT KINASE 2; CHAIN: NULL;	PROTEIN KINASE CDK2; TRANSFERASE, SERINE/THREONINE PROTEIN KINASE,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										ATP-BINDING, 2 CELL CYCLE, CELL DIVISION, MITOSIS, PHOSPHORYLATION
710	1ian		8	377	3.6e-43			108.46	P38 MAP KINASE; CHAIN: NULL;	SERINE/THREONINE-PROTEIN KINASE CSBP, RK, P38; PROTEIN SER/THR-KINASE, SERINE/THREONINE-PROTEIN KINASE
710	1ir3	A	9	312	1.6e-31			91.02	INSULIN RECEPTOR; CHAIN: A; PEPTIDE SUBSTRATE; CHAIN: B;	COMPLEX (TRANSFERASE/SUBSTRATE) TYROSINE KINASE, SIGNAL TRANSDUCTION, PHOSPHOTRANSFERASE, 2 COMPLEX (KINASE/PEPTIDE SUBSTRATE/ATP ANALOG), ENZYME, 3 COMPLEX (TRANSFERASE/SUBSTRATE)
710	1jnk		20	299	3.6e-45	0.45	1.00		C-JUN N-TERMINAL KINASE; CHAIN: NULL;	TRANSFERASE JNK3; TRANSFERASE, JNK3 MAP KINASE, SERINE/THREONINE PROTEIN 2 KINASE
710	1jnk		4	360	3.6e-45			109.60	C-JUN N-TERMINAL KINASE; CHAIN: NULL;	TRANSFERASE JNK3; TRANSFERASE, JNK3 MAP KINASE, SERINE/THREONINE PROTEIN 2 KINASE
710	1koa		1	414	6e-68			136.94	TWITCHIN; CHAIN: NULL;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
710	1koa		21	339	6e-68	0.33	1.00		TWITCHIN; CHAIN: NULL;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
710	1koa		9	282	9e-70	0.09	1.00		TWITCHIN; CHAIN: NULL;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
710	1kob	A	5	345	1.4e-71			142.96	TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
710	1kob	A	9	294	1.4e-71	0.45	1.00		TWITCHIN; CHAIN: A, B;	KINASE KINASE, TWITCHIN, INTRASTERIC REGULATION
710	1p38		20	311	1.4e-50	0.34	1.00		MAP KINASE P38; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE; TRANSFERASE, MAP KINASE, SERINE/THREONINE-PROTEIN KINASE, 2 P38
710	1p38		7	348	1.4e-50			119.99	MAP KINASE P38; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE; TRANSFERASE, MAP KINASE, SERINE/THREONINE-PROTEIN KINASE, 2 P38
710	1phk		20	287	7.2e-86			132.07	PHOSPHORYLASE KINASE;	KINASE RABBIT MUSCLE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: NULL;	PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE-PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
710	1phk		21	284	7.2e-86	0.49	1.00		PHOSPHORYLASE KINASE; CHAIN: NULL;	KINASE RABBIT MUSCLE PHOSPHORYLASE KINASE; GLYCOGEN METABOLISM, TRANSFERASE, SERINE/THREONINE-PROTEIN, 2 KINASE, ATP-BINDING, CALMODULIN-BINDING
710	1pme		17	331	7.2e-43			103.32	ERK2; CHAIN: NULL;	TRANSFERASE MAP KINASE, SERINE/THREONINE PROTEIN KINASE, TRANSFERASE
710	1tki	A	17	354	1.4e-56			108.54	TITIN; CHAIN: A, B;	SERINE KINASE SERINE KINASE, TITIN, MUSCLE, AUTOINHIBITION
710	1tki	A	21	284	1.4e-56	0.44	1.00		TITIN; CHAIN: A, B;	SERINE KINASE SERINE KINASE, TITIN, MUSCLE, AUTOINHIBITION
710	3erk		1	342	7.2e-45			119.69	EXTRACELLULAR REGULATED KINASE 2; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE, MAP 2, ERK2; TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, MAP KINASE, 2 ERK2
710	3erk		22	301	7.2e-45	0.44	1.00		EXTRACELLULAR REGULATED KINASE 2; CHAIN: NULL;	TRANSFERASE MITOGEN ACTIVATED PROTEIN KINASE, MAP 2, ERK2; TRANSFERASE, SERINE/THREONINE-PROTEIN KINASE, MAP KINASE, 2 ERK2
721	1buo	A	171	297	4e-21			59.40	PROMYELOCYTIC LEUKEMIA ZINC FINGER PROTEIN PLZF; CHAIN: A;	GENE REGULATION POZ DOMAIN; PROTEIN-PROTEIN INTERACTION DOMAIN, TRANSCRIPTIONAL 2 REPRESSOR, ZINC-FINGER PROTEIN, X-RAY CRYSTALLOGRAPHY, 3 PROTEIN STRUCTURE, PROMYELOCYTIC LEUKEMIA, GENE REGULATION
721	1buo	A	172	294	1.8e-15	0.74	1.00		PROMYELOCYTIC LEUKEMIA ZINC FINGER PROTEIN PLZF; CHAIN: A;	GENE REGULATION POZ DOMAIN; PROTEIN-PROTEIN INTERACTION DOMAIN, TRANSCRIPTIONAL 2 REPRESSOR, ZINC-FINGER PROTEIN, X-RAY CRYSTALLOGRAPHY, 3 PROTEIN STRUCTURE, PROMYELOCYTIC

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										LEUKEMIA, GENE REGULATION
721	1buo	A	185	293	4e-21	0.34	0.94		PROMYELOCYTIC LEUKEMIA ZINC FINGER PROTEIN PLZF; CHAIN: A;	GENE REGULATION POZ DOMAIN; PROTEIN-PROTEIN INTERACTION DOMAIN, TRANSCRIPTIONAL 2 REPRESSOR, ZINC-FINGER PROTEIN, X-RAY CRYSTALLOGRAPHY, 3 PROTEIN STRUCTURE, PROMYELOCYTIC LEUKEMIA, GENE REGULATION
721	1ca9	A	19	162	3.6e-22	0.37	0.88		TNF RECEPTOR ASSOCIATED FACTOR 2; CHAIN: A, B, C, D, E, F; TNF-R2; CHAIN: G, H;	TNF SIGNALING TRAF2; TNF SIGNALING, TRAF, ADAPTER PROTEIN, CELL SURVIVAL
721	1czy	A	19	162	5.4e-22	0.34	0.59		TUMOR NECROSIS FACTOR RECEPTOR ASSOCIATED CHAIN: A, B, C; LATENT MEMBRANE PROTEIN 1; CHAIN: D, E;	APOPTOSIS TRAF2; LMP1; BETA SANDWICH, PROTEIN-PEPTIDE COMPLEX, SIGNALING PROTEIN
721	1czy	A	20	164	4e-26	0.60	0.53		TUMOR NECROSIS FACTOR RECEPTOR ASSOCIATED CHAIN: A, B, C; LATENT MEMBRANE PROTEIN 1; CHAIN: D, E;	APOPTOSIS TRAF2; LMP1; BETA SANDWICH, PROTEIN-PEPTIDE COMPLEX, SIGNALING PROTEIN
721	1czz	A	19	162	3.6e-21	0.51	0.88		TUMOR NECROSIS FACTOR RECEPTOR ASSOCIATED CHAIN: A, B, C; CD 40 PEPTIDE; CHAIN: D, E;	APOPTOSIS TRAF2; CD40; B-SANDWICH, PROTEIN-PEPTIDE COMPLEX
721	1czz	A	20	164	4e-28	0.40	0.86		TUMOR NECROSIS FACTOR RECEPTOR ASSOCIATED CHAIN: A, B, C; CD 40 PEPTIDE; CHAIN: D, E;	APOPTOSIS TRAF2; CD40; B-SANDWICH, PROTEIN-PEPTIDE COMPLEX
721	1flk	A	1	162	1.3e-20	0.23	0.28		TNF RECEPTOR ASSOCIATED FACTOR 3; CHAIN: A, B;	APOPTOSIS TNF SIGNALING, TRAF3, CD40-BINDING PROTEIN
721	1qsc	A	19	162	3.6e-22	0.33	0.84		TNF RECEPTOR ASSOCIATED FACTOR 2; CHAIN: A, B, C; CD40 RECEPTOR; CHAIN: D, E, F;	SIGNALING PROTEIN TRAF2; TNF SIGNALING, TRAF, CD40 RECEPTOR, ADAPTER PROTEIN, CELL 2 SURVIVAL, COILED COIL, SIGNALING PROTEIN
723	1eqz	B	36	87	1.3e-25	-0.51	1.00		HISTONE H2A; CHAIN: A, E; HISTONE H2B; CHAIN: B, F; HISTONE H3; CHAIN: C, G; HISTONE H4; CHAIN: D, H; 146	STRUCTURAL PROTEIN/DNA NUCLEOSOME, NUCLEOSOME CORE PARTICLE, HISTONE, MICROGRAVITY 2 HISTONE OCTAMER, DNA PALINDROME,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									NUCLEOTIDES LONG DNA; CHAIN: I, J;	DNA PROTEIN COMPLEX, 3 CHROMATIN, CHROMOSOMAL PROTEIN, HISTONE FOLD, BENT DNA
723	1f66	D	36	87	3.6e-25	-0.24	0.99		HISTONE H3; CHAIN: A, E; HISTONE H4; CHAIN: B, F; HISTONE H2A.Z; CHAIN: C, G; HISTONE H2B; CHAIN: D, H; PALINDROMIC 146 BASE PAIR DNA FRAGMENT; CHAIN: I, J;	STRUCTURAL PROTEIN/DNA NUCLEOSOME, CHROMATIN, HISTONE, HISTONE VARIANT, PROTEIN 2 DNA INTERACTION, NUCLEOPROTEIN, SUPERCOILED DNA,COMPLEX 3 (NUCLEOSOME CORE/DNA)
723	1hio	B	37	87	9e-23	-0.32	0.96		HISTONE H2A; CHAIN: A; HISTONE H2B; CHAIN: B; HISTONE H3; CHAIN: C; HISTONE H4; CHAIN: D;	CHROMOSOMAL PROTEIN HISTONE, CHROMOSOMAL PROTEIN, NUCLEOSOME CORE
724	1b8q	A	11	99	1.2e-15	0.35	0.93		NEURONAL NITRIC OXIDE SYNTHASE; CHAIN: A; HEPTAPEPTIDE; CHAIN: B;	OXIDOREDUCTASE PDZ DOMAIN, NNOS, NITRIC OXIDE SYNTHASE
724	1be9	A	5	99	1.8e-12	0.68	0.99		PSD-95; CHAIN: A; CRIPT; CHAIN: B;	PEPTIDE RECOGNITION PEPTIDE RECOGNITION, PROTEIN LOCALIZATION
724	1kwa	A	5	85	6e-16	0.74	0.98		HCASK/LIN-2 PROTEIN; CHAIN: A, B;	KINASE HCASK, GLGF REPEAT, DHR; PDZ DOMAIN, NEUREXIN, SYNDECAN, RECEPTOR CLUSTERING, KINASE
724	1pdr		4	91	9e-12	0.34	1.00		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
724	1pdr		5	86	1e-13	0.76	1.00		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
724	1qau	A	2	79	5.4e-07	0.39	0.92		NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: A;	OXIDOREDUCTASE BETA-FINGER
724	1qav	A	3	82	3.6e-12	0.55	1.00		ALPHA-1 SYNTROPHIN (RESIDUES 77-171); CHAIN: A; NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: B;	MEMBRANE PROTEIN/OXIDOREDUCTASE BETA-FINGER, HETERODIMER
724	1qlc	A	1	80	1.1e-09	0.51	0.82		POSTSYNAPTIC DENSITY PROTEIN 95; CHAIN: A;	PEPTIDE RECOGNITION PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										SYNTHASE, NMDA RECEPTOR 2 BINDING
724	1qlc	A	3	85	4e-16	0.55	0.99		POSTSYNAPTIC DENSITY PROTEIN 95; CHAIN: A;	PEPTIDE RECOGNITION PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE SYNTHASE, NMDA RECEPTOR 2 BINDING
724	3pdz	A	8	85	5.4e-09	0.83	0.93		TYROSINE PHOSPHATASE (PTP-BAS, TYPE 1); CHAIN: A;	HYDROLASE PDZ DOMAIN, HUMAN PHOSPHATASE, HPTP1E, PTP-BAS, SPECIFICITY 2 OF BINDING
724	1b8q	A	11	99	1.2e-15	0.35	0.93		NEURONAL NITRIC OXIDE SYNTHASE; CHAIN: A; HEPTAPEPTIDE; CHAIN: B;	OXIDOREDUCTASE PDZ DOMAIN, NNOS, NITRIC OXIDE SYNTHASE
724	1b8q	A	7	114	3.6e-06	0.00	0.71		NEURONAL NITRIC OXIDE SYNTHASE; CHAIN: A; HEPTAPEPTIDE; CHAIN: B;	OXIDOREDUCTASE PDZ DOMAIN, NNOS, NITRIC OXIDE SYNTHASE
724	1be9	A	4	84	5.4e-11	0.74	1.00		PSD-95; CHAIN: A; CRIPT; CHAIN: B;	PEPTIDE RECOGNITION PEPTIDE RECOGNITION, PROTEIN LOCALIZATION
724	1i16		4	76	3.6e-06	0.64	0.92		INTERLEUKIN 16; CHAIN: NULL;	CYTOKINE LCF; CYTOKINE, LYMPHOCYTE CHEMOATTRACTANT FACTOR, PDZ DOMAIN
724	1kwa	A	5	85	6e-16	0.74	0.98		HCASK/LIN-2 PROTEIN; CHAIN: A, B;	KINASE HCASK, GLGF REPEAT, DHR; PDZ DOMAIN, NEUREXIN, SYNDECAN, RECEPTOR CLUSTERING, KINASE
724	1pdr		5	84	3.6e-10	0.51	1.00		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
724	1pdr		5	86	1e-13	0.76	1.00		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
724	1qau	A	7	102	3.6e-05	0.24	0.15		NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: A;	OXIDOREDUCTASE BETA-FINGER
724	1qav	A	4	82	9e-10	0.80	1.00		ALPHA-1 SYNTROPHIN (RESIDUES 77-171); CHAIN: A; NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: B;	MEMBRANE PROTEIN/OXIDOREDUCTASE BETA-FINGER, HETERODIMER
724	1qlc	A	3	85	4e-16	0.55	0.99		POSTSYNAPTIC DENSITY PROTEIN 95; CHAIN: A;	PEPTIDE RECOGNITION PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE SYNTHASE, NMDA RECEPTOR 2 BINDING

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
724	3pdz	A	11	86	7.2e-08	0.62	0.83		TYROSINE PHOSPHATASE (PTP-BAS, TYPE 1); CHAIN: A;	HYDROLASE PDZ DOMAIN, HUMAN PHOSPHATASE, HPTP1E, PTP-BAS, SPECIFICITY 2 OF BINDING
725	1b8q	A	11	99	1.2e-15	0.35	0.93		NEURONAL NITRIC OXIDE SYNTHASE; CHAIN: A; HEPTAPEPTIDE; CHAIN: B;	OXIDOREDUCTASE PDZ DOMAIN, NNOS, NITRIC OXIDE SYNTHASE
725	1be9	A	5	99	1.8e-12	0.68	0.99		PSD-95; CHAIN: A; CRIPT; CHAIN: B;	PEPTIDE RECOGNITION PEPTIDE RECOGNITION, PROTEIN LOCALIZATION
725	1kwa	A	5	85	6e-16	0.74	0.98		HCASK/LIN-2 PROTEIN; CHAIN: A, B;	KINASE HCASK, GLGF REPEAT, DHR; PDZ DOMAIN, NEUREXIN, SYNDECAN, RECEPTOR CLUSTERING, KINASE
725	1pdr		4	91	9e-12	0.34	1.00		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
725	1pdr		5	86	1e-13	0.76	1.00		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
725	1qau	A	2	79	5.4e-07	0.39	0.92		NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: A;	OXIDOREDUCTASE BETA-FINGER
725	1qav	A	3	82	3.6e-12	0.55	1.00		ALPHA-1 SYNTROPHIN (RESIDUES 77-171); CHAIN: A; NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: B;	MEMBRANE PROTEIN/OXIDOREDUCTASE BETA-FINGER, HETERODIMER
725	1qlc	A	1	80	1.1e-09	0.51	0.82		POSTSYNAPTIC DENSITY PROTEIN 95; CHAIN: A;	PEPTIDE RECOGNITION PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE SYNTHASE, NMDA RECEPTOR 2 BINDING
725	1qlc	A	3	85	4e-16	0.55	0.99		POSTSYNAPTIC DENSITY PROTEIN 95; CHAIN: A;	PEPTIDE RECOGNITION PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE SYNTHASE, NMDA RECEPTOR 2 BINDING
725	3pdz	A	8	85	5.4e-09	0.83	0.93		TYROSINE PHOSPHATASE (PTP-BAS, TYPE 1); CHAIN: A;	HYDROLASE PDZ DOMAIN, HUMAN PHOSPHATASE, HPTP1E, PTP-BAS, SPECIFICITY 2 OF BINDING
725	1b8q	A	11	99	1.2e-15	0.35	0.93		NEURONAL NITRIC OXIDE SYNTHASE; CHAIN: A; HEPTAPEPTIDE; CHAIN: B;	OXIDOREDUCTASE PDZ DOMAIN, NNOS, NITRIC OXIDE SYNTHASE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
725	1b8q	A	7	114	3.6e-06	0.00	0.71		NEURONAL NITRIC OXIDE SYNTHASE; CHAIN: A; HEPTAPEPTIDE; CHAIN: B;	OXIDOREDUCTASE PDZ DOMAIN, NNOS, NITRIC OXIDE SYNTHASE
725	1be9	A	4	84	5.4e-11	0.74	1.00		PSD-95; CHAIN: A; CRIPT; CHAIN: B;	PEPTIDE RECOGNITION PEPTIDE RECOGNITION, PROTEIN LOCALIZATION
725	1i16		4	76	3.6e-06	0.64	0.92		INTERLEUKIN 16; CHAIN: NULL;	CYTOKINE LCF; CYTOKINE, LYMPHOCYTE CHEMOATTRACTANT FACTOR, PDZ DOMAIN
725	1kwa	A	5	85	6e-16	0.74	0.98		HCASK/LIN-2 PROTEIN; CHAIN: A, B;	KINASE HCASK, GLGF REPEAT, DHR; PDZ DOMAIN, NEUREXIN, SYNDECAN, RECEPTOR CLUSTERING, KINASE
725	1pdr		5	84	3.6e-10	0.51	1.00		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
725	1pdr		5	86	1e-13	0.76	1.00		HUMAN DISCS LARGE PROTEIN; CHAIN: NULL;	SIGNAL TRANSDUCTION HDLG, DHR3 DOMAIN; SIGNAL TRANSDUCTION, SH3 DOMAIN, REPEAT
725	1qau	A	7	102	3.6e-05	0.24	0.15		NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: A;	OXIDOREDUCTASE BETA-FINGER
725	1qav	A	4	82	9e-10	0.80	1.00		ALPHA-1 SYNTROPHIN (RESIDUES 77-171); CHAIN: A; NEURONAL NITRIC OXIDE SYNTHASE (RESIDUES 1-130); CHAIN: B;	MEMBRANE PROTEIN/OXIDOREDUCTASE BETA-FINGER, HETERODIMER
725	1qlc	A	3	85	4e-16	0.55	0.99		POSTSYNAPTIC DENSITY PROTEIN 95; CHAIN: A;	PEPTIDE RECOGNITION PSD-95; PDZ DOMAIN, NEURONAL NITRIC OXIDE SYNTHASE, NMDA RECEPTOR 2 BINDING
725	3pdz	A	11	86	7.2e-08	0.62	0.83		TYROSINE PHOSPHATASE (PTP-BAS, TYPE 1); CHAIN: A;	HYDROLASE PDZ DOMAIN, HUMAN PHOSPHATASE, HPTP1E, PTP-BAS, SPECIFICITY 2 OF BINDING
726	1axi	B	44	226	4e-05			59.70	GROWTH HORMONE; CHAIN: A; GROWTH HORMONE RECEPTOR; CHAIN: B;	COMPLEX (HORMONE/RECEPTOR) HGH; HGHBP; COMPLEX (HORMONE/RECEPTOR)
726	1bj8		125	222	8e-14	0.22	0.40		GP130; CHAIN: NULL;	RECEPTOR RECEPTOR, SIGNAL TRANSDUCER OF IL-6 TYPE CYTOKINES, THIRD 2 N-TERMINAL DOMAIN,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										TRANSMEMBRANE, GLYCOPROTEIN
726	1bp3	B	25	225	8e-10			57.30	GROWTH HORMONE; CHAIN: A; PROLACTIN RECEPTOR; CHAIN: B;	HORMONE/GROWTH FACTOR HORMONE, RECEPTOR, HORMONE/GROWTH FACTOR
726	1bpv		128	227	8e-13	0.47	0.41		TITIN; CHAIN: NULL;	CONNECTIN A71, CONNECTIN; TITIN, CONNECTIN, FIBRONECTIN TYPE III
726	1cfb		127	222	1.6e-10	0.06	0.42		NEURAL ADHESION MOLECULE DROSOPHILA NEUROGLIAN (CHYMOTRYPTIC FRAGMENT CONTAINING THE 1CFB 3 TWO AMINO PROXIMAL FIBRONECTIN TYPE III REPEATS 1CFB 4 (RESIDUES 610 - 814)) 1CFB 5	
726	1fnf		128	222	4e-12	0.10	0.42		FIBRONECTIN; 1FNF 6 CHAIN: NULL; 1FNF 7	CELL ADHESION PROTEIN RGD, EXTRACELLULAR MATRIX 1FNF 18
726	1mfn		128	223	1.6e-11	-0.01	0.11		FIBRONECTIN; CHAIN: NULL;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN, RGD, EXTRACELLULAR MATRIX, 2 HEPARIN-BINDING, GLYCOPROTEIN
726	1mfn		131	241	1.4e-10	0.18	0.05		FIBRONECTIN; CHAIN: NULL;	CELL ADHESION PROTEIN CELL ADHESION PROTEIN, RGD, EXTRACELLULAR MATRIX, 2 HEPARIN-BINDING, GLYCOPROTEIN
726	2fnb	A	131	222	2e-12	0.19	0.34		FIBRONECTIN; CHAIN: A;	PROTEIN BINDING ED-B, FIBRONECTIN, TYPEIII DOMAIN, ANGIOGENESIS, PROTEIN 2 BINDING
726	2hft		128	234	1.2e-10	0.05	-0.11		HUMAN TISSUE FACTOR; 2HFT 4 CHAIN: NULL; 2HFT 5	COAGULATION FACTOR
727	1am4	D	20	191	1.6e-48			63.49	P50-RHOGAP; CHAIN: A, B, C; CDC42HS; CHAIN: D, E, F;	COMPLEX (GTPASE-ACTIVATING/GTP-BINDING) COMPLEX (GTPASE-ACTIVATING/GTP-BINDING), GTPASE ACTIVATION
727	1byu	A	20	232	5.4e-37			69.00	GTP-BINDING PROTEIN RAN; CHAIN: A, B;	TRANSPORT PROTEIN TC4; GTPASE, NUCLEAR TRANSPORT, TRANSPORT PROTEIN
727	1byu	B	14	237	5.4e-37			67.40	GTP-BINDING PROTEIN RAN;	TRANSPORT PROTEIN TC4; GTPASE,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: A, B;	NUCLEAR TRANSPORT, TRANSPORT PROTEIN
727	1cly	A	20	193	5.4e-66			101.19	RAS-RELATED PROTEIN RAP-1A; CHAIN: A; PROTO-ONKOGENE SERINE/THREONINE PROTEIN KINASE CHAIN: B;	SIGNALING PROTEIN GTP-BINDING PROTEINS, PROTEIN-PROTEIN COMPLEX, EFFECTORS
727	1ctq	A	20	194	9e-67			109.09	TRANSFORMING PROTEIN P21/H-RAS-1; CHAIN: A;	SIGNALING PROTEIN G PROTEIN, GTP HYDROLYSIS, KINETIC CRYSTALLOGRAPHY, 2 SIGNALING PROTEIN
727	1cxz	A	17	194	7.2e-54			67.75	HIS-TAGGED TRANSFORMING PROTEIN RHOA(0-181); CHAIN: A; PKN; CHAIN: B;	SIGNALING PROTEIN PROTEIN-PROTEIN COMPLEX, ANTIPARALLEL COILED-COIL
727	1ibr	A	21	198	1.8e-36			67.32	RAN; CHAIN: A, C; IMPORTIN BETA SUBUNIT; CHAIN: B, D;	SMALL GTPASE KARYOPHERIN BETA, P95 SMALL GTPASE, NUCLEAR TRANSPORT RECEPTOR
727	1kao		20	194	1.1e-62			115.88	RAP2A; CHAIN: NULL;	GTP-BINDING PROTEIN GTP-BINDING PROTEIN, SMALL G PROTEIN, RAP2, GDP, RAS
727	1mh1		19	196	3.6e-56			75.70	RAC1; CHAIN: NULL;	GTP-BINDING GTP-BINDING, GTPASE, SMALL G-PROTEIN, RHO FAMILY, RAS SUPER 2 FAMILY
727	1plj		23	193	1.4e-49			58.99	ONCOGENE PROTEIN C-H-RAS P21 PROTEIN MUTANT WITH GLY 12 REPLACED BY PRO 1PLJ 3 (G12P) COMPLEXED WITH P3-1-(2-NITROPHENYL)ETHYL- 1PLJ 4 GUANOSINE-5'-(B,G-IMIDO)-TRIPHOSPHATE 1PLJ 5	
727	1rrp	C	21	215	1.8e-36			70.51	RAN; CHAIN: A, C; NUCLEAR PORE COMPLEX PROTEIN NUP358; CHAIN: B, D;	COMPLEX (SMALL GTPASE/NUCLEAR PROTEIN) COMPLEX (SMALL GTPASE/NUCLEAR PROTEIN), SMALL GTPASE, 2 NUCLEAR TRANSPORT
727	1tx4	B	20	191	1.6e-50			62.17	P50-RHOGAP; CHAIN: A; TRANSFORMING PROTEIN RHOA; CHAIN: B;	COMPLEX(GTPASE ACTIVATN/PROTO-ONCOGENE) GTPASE-ACTIVATING PROTEIN RHOGAP; COMPLEX (GTPASE ACTIVATION/PROTO-ONCOGENE), GTPASE, 2 TRANSITION STATE, GAP

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
727	1zbd	A	13	199	5.4e-63			78.74	RAB-3A; CHAIN: A; RABPHILIN-3A; CHAIN: B;	COMPLEX (GTP-BINDING/EFFECTOR) RAS-RELATED PROTEIN RAB3A; COMPLEX (GTP-BINDING/EFFECTOR), G PROTEIN, EFFECTOR, RABCDR, 2 SYNAPTIC EXOCYTOSIS, RAB PROTEIN, RAB3A, RABPHILIN
727	2ngr	A	20	208	1.6e-50			65.77	GTP BINDING PROTEIN (G25K); CHAIN: A; GTPASE ACTIVATING PROTEIN (RHG); CHAIN: B;	HYDROLASE CDC42/CDC42GAP; CDC42/CDC42GAP; TRANSITION STATE, G-PROTEIN, GAP, CDC42, ALF3., HYDROLASE
727	3rab	A	14	194	7.2e-63			90.71	RAB3A; CHAIN: A;	HYDROLASE G PROTEIN, VESICULAR TRAFFICKING, GTP HYDROLYSIS, RAB 2 PROTEIN, NEUROTRANSMITTER RELEASE, HYDROLASE
731	1tgx	A	55	98	0.0031	-0.49	0.01		CYTOTOXIN TOXIN GAMMA (CARDIOTOXIN) 1TGX 3	
731	2crs		55	98	0.0023	-0.25	0.00		CARDIOTOXIN CARDIOTOXIN III (NMR, 13 STRUCTURES) 2CRS 3	
732	1b0w	A	20	130	3.6e-47			52.89	BENCE-JONES KAPPA I PROTEIN BRE; CHAIN: A, B, C;	IMMUNE SYSTEM BENCE-JONES; IMMUNOGLOBULIN, AMYLOID, IMMUNE SYSTEM
732	1b6d	A	20	126	1.1e-49	-0.03	0.98		IMMUNOGLOBULIN; CHAIN: A, B;	IMMUNOGLOBULIN IMMUNOGLOBULIN, KAPPA LIGHT-CHAIN DIMER HEADER
732	1bj1	L	20	126	3.6e-50	0.13	0.94		FAB FRAGMENT; CHAIN: L, H, J, K; VASCULAR ENDOTHELIAL GROWTH FACTOR; CHAIN: V, W;	COMPLEX (ANTIBODY/ANTIGEN) FAB-12; VEGF; COMPLEX (ANTIBODY/ANTIGEN), ANGIOGENIC FACTOR
732	1bvk	A	20	130	5.4e-47			51.11	HULYS11; CHAIN: A, B, D, E; LYSOZYME; CHAIN: C, F;	COMPLEX (HUMANIZED ANTIBODY/HYDROLASE) MURAMIDASE; HUMANIZED ANTIBODY, ANTIBODY COMPLEX, FV, ANTI-LYSOZYME, 2 COMPLEX (HUMANIZED ANTIBODY/HYDROLASE)
732	1bww	A	18	129	1.8e-49			52.07	IG KAPPA CHAIN V-I REGION REI; CHAIN: A, B;	IMMUNE SYSTEM REIV, STABILIZED IMMUNOGLOBULIN FRAGMENT, BENCE-JONES 2 PROTEIN, IMMUNE SYSTEM
732	1bww	A	20	127	1.8e-49	0.17	1.00		IG KAPPA CHAIN V-I REGION	IMMUNE SYSTEM REIV, STABILIZED

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									REI; CHAIN: A, B;	IMMUNOGLOBULIN FRAGMENT, BENCE-JONES 2 PROTEIN, IMMUNE SYSTEM
732	1dee	A	20	126	3.6e-52	0.19	1.00		IGM RF 2A2; CHAIN: A, C, E; IGM RF 2A2; CHAIN: B, D, F; IMMUNOGLOBULIN G BINDING PROTEIN A; CHAIN: G, H;	IMMUNE SYSTEM FAB-IBP COMPLEX CRYSTAL STRUCTURE 2.7A RESOLUTION BINDING 2 OUTSIDE THE ANTIGEN COMBINING SITE SUPERANTIGEN FAB VH3 3 SPECIFICITY
732	1fgv	L	20	126	7.2e-51	0.27	0.98		IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 1FGV 3 ANTIBODY 'H52' (HUH52-AA FV) 1FGV 4	
732	1fgv	L	20	129	7.2e-51			54.34	IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 1FGV 3 ANTIBODY 'H52' (HUH52-AA FV) 1FGV 4	
732	1fvc	A	20	126	1.3e-48	0.31	0.98		IMMUNOGLOBULIN FV FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 8 1FVC 3	
732	1fvc	A	20	130	1.3e-48			50.75	IMMUNOGLOBULIN FV FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 8 1FVC 3	
732	1fvd	A	20	126	5.4e-49	0.06	0.95		IMMUNOGLOBULIN FAB FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 4 1FVD 3	
732	1igm	L	20	126	3.6e-48	-0.16	0.89		IMMUNOGLOBULIN IMMUNOGLOBULIN M (IG-M) FV FRAGMENT 1IGM 3	
732	1igm	L	20	130	3.6e-48			50.56	IMMUNOGLOBULIN IMMUNOGLOBULIN M (IG-M) FV FRAGMENT 1IGM 3	
732	1nmb	L	20	130	1.8e-42			52.25	N9 NEURAMINIDASE; 1NMB 4 CHAIN: N; 1NMB 5 FAB NC10; 1NMB 9 CHAIN: L, H; 1NMB 10	COMPLEX (HYDROLASE/IMMUNOGLOBULIN)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
732	1tcr	A	21	128	1.4e-41			64.90	ALPHA, BETA T-CELL RECEPTOR CHAIN: A, B;	RECEPTOR TCR; T-CELL, RECEPTOR, TRANSMEMBRANE, GLYCOPROTEIN, SIGNAL
732	1wtl	A	20	126	7.2e-48	0.20	0.94		IMMUNOGLOBULIN WAT, A VARIABLE DOMAIN FROM IMMUNOGLOBULIN LIGHT-CHAIN 1WTL 3 (BENCE-JONES PROTEIN) 1WTL 4	
732	1wtl	A	20	129	7.2e-48			51.95	IMMUNOGLOBULIN WAT, A VARIABLE DOMAIN FROM IMMUNOGLOBULIN LIGHT-CHAIN 1WTL 3 (BENCE-JONES PROTEIN) 1WTL 4	
732	2fgw	L	20	126	7.2e-51	-0.17	0.99		IMMUNOGLOBULIN FAB FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 2FGW 3 ANTIBODY 'H52' (HUH52-OZ FAB) 2FGW 4	
735	1ez3	A	24	145	6e-09	0.24	-0.05		SYNTAXIN-1A; CHAIN: A, B, C;	ENDOCYTOSIS/EXOCYTOSIS SYNAPTOTAGMIN ASSOCIATED 35 KDA PROTEIN, P35A, THREE HELIX BUNDLE
735	1fio	A	22	182	2e-05	-0.11	0.06		SSO1 PROTEIN; CHAIN: A;	MEMBRANE PROTEIN FOUR HELIX BUNDLE, ALPHA HELIX
735	1ses	A	23	87	3.6e-06	0.43	0.01		LIGASE(SYNTHETASE) SERYL-TRNA SYNTHETASE (E.C.6.1.1.11) (SERINE-TRNA LIGASE) 1SES 3 COMPLEXED WITH SERYL-HYDROXAMATE-AMP 1SES 4	
738	1aj4		19	145	3.6e-37	-0.16	0.99		TROPONIN C; CHAIN: NULL;	MUSCLE PROTEIN CTNC; CARDIAC, MUSCLE PROTEIN, REGULATORY, CALCIUM BINDING
738	1ak8		17	92	1.3e-30			52.90	CALMODULIN; CHAIN: NULL;	CALCIUM-BINDING PROTEIN CALMODULIN CERIUM TR1C-DOMAIN, RESIDUES 1 - 75; CERIUM-LOADED, CALCIUM-BINDING PROTEIN
738	1ak8		17	93	1.3e-30	0.30	0.98		CALMODULIN; CHAIN: NULL;	CALCIUM-BINDING PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										CALMODULIN CERIUM TRIC-DOMAIN, RESIDUES 1 - 75; CERIUM-LOADED, CALCIUM-BINDING PROTEIN
738	1avs	A	16	94	3.6e-26			50.15	TROPONIN C; CHAIN: A, B;	MUSCLE CONTRACTION MUSCLE CONTRACTION, CALCIUM-ACTIVATED, TROPONIN, E-F HAND 2 CALCIUM-BINDING PROTEIN
738	1cdm	A	21	144	1.8e-45	0.19	0.96		CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING DOMAIN OF 1CDM 3 CALMODULIN-DEPENDENT PROTEIN KINASE II 1CDM 4	
738	1cdm	A	21	145	1.8e-45			61.02	CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING DOMAIN OF 1CDM 3 CALMODULIN-DEPENDENT PROTEIN KINASE II 1CDM 4	
738	1cll		21	144	9e-50	0.07	0.96		CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) 1CLL 3	
738	1cll		21	145	9e-50			59.05	CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) 1CLL 3	
738	1dtl	A	19	145	3.6e-36	-0.01	0.89		CARDIAC TROPONIN C; CHAIN: A;	STRUCTURAL PROTEIN HELIX-TURN-HELIX
738	1exr	A	19	144	1.8e-47	0.17	0.96		CALMODULIN; CHAIN: A;	METAL TRANSPORT CALMODULIN, HIGH RESOLUTION, DISORDER
738	1tcf		16	144	1.4e-39			57.73	TROPONIN C; CHAIN: NULL;	CALCIUM-REGULATED MUSCLE CONTRACTION MUSCLE CONTRACTION, CALCIUM-BINDING, TROPONIN, E-F HAND, 2 OPEN CONFORMATION REGULATORY DOMAIN, CALCIUM-REGULATED 3 MUSCLE CONTRACTION
738	1tcf		21	145	1.4e-39	0.16	0.96		TROPONIN C; CHAIN: NULL;	CALCIUM-REGULATED MUSCLE CONTRACTION MUSCLE CONTRACTION, CALCIUM-BINDING, TROPONIN, E-F

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										HAND, 2 OPEN CONFORMATION REGULATORY DOMAIN, CALCIUM-REGULATED 3 MUSCLE CONTRACTION
738	1tnx		16	144	3.6e-36			55.28	TROPONIN C; 1TNX 4 CHAIN: NULL; 1TNX 5	CALCIUM-BINDING PROTEIN EF-HAND 1TNX 14
738	1tnx		21	145	3.6e-36	-0.26	0.83		TROPONIN C; 1TNX 4 CHAIN: NULL; 1TNX 5	CALCIUM-BINDING PROTEIN EF-HAND 1TNX 14
738	1top		21	145	7.2e-40	0.40	1.00		CONTRACTILE SYSTEM PROTEIN TROPONIN C 1TOP 3	
738	1top		3	144	7.2e-40			53.77	CONTRACTILE SYSTEM PROTEIN TROPONIN C 1TOP 3	
738	1vrk	A	18	144	3.6e-49	0.15	0.99		CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALLING, 2 COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
738	1vrk	A	18	144	3.6e-49			58.84	CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALLING, 2 COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
739	1aj4		11	154	7.2e-37			58.71	TROPONIN C; CHAIN: NULL;	MUSCLE PROTEIN CTNC; CARDIAC, MUSCLE PROTEIN, REGULATORY, CALCIUM BINDING
739	1aj4		19	154	7.2e-37	-0.00	1.00		TROPONIN C; CHAIN: NULL;	MUSCLE PROTEIN CTNC; CARDIAC, MUSCLE PROTEIN, REGULATORY, CALCIUM BINDING
739	1ak8		17	92	3.6e-30			52.74	CALMODULIN; CHAIN: NULL;	CALCIUM-BINDING PROTEIN CALMODULIN CERIUM TR1C-DOMAIN, RESIDUES 1 - 75; CERIUM-LOADED, CALCIUM-BINDING PROTEIN
739	1brl	B	21	144	3.6e-33	0.07	1.00		MYOSIN; CHAIN: A, B, C, D, E, F, G, H;	MUSCLE PROTEIN MDE; MUSCLE PROTEIN
739	1cdm	A	21	144	7.2e-47	0.25	1.00		CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING DOMAIN OF 1CDM 3 CALMODULIN-DEPENDENT PROTEIN KINASE II 1CDM 4	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
739	1cdm	A	21	147	7.2e-47			58.78	CALCIUM-BINDING PROTEIN CALMODULIN COMPLEXED WITH CALMODULIN-BINDING DOMAIN OF 1CDM 3 CALMODULIN-DEPENDENT PROTEIN KINASE II 1CDM 4	
739	1cll		21	144	5.4e-52	0.04	1.00		CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) 1CLL 3	
739	1cll		21	154	5.4e-52			67.13	CALCIUM-BINDING PROTEIN CALMODULIN (VERTEBRATE) 1CLL 3	
739	1dtl	A	19	154	1.3e-32	-0.08	0.68		CARDIAC TROPONIN C; CHAIN: A;	STRUCTURAL PROTEIN HELIX-TURN-HELIX
739	1exr	A	19	144	1.6e-49	-0.08	1.00		CALMODULIN; CHAIN: A;	METAL TRANSPORT CALMODULIN, HIGH RESOLUTION, DISORDER
739	1tcf		16	154	1.8e-40			65.99	TROPONIN C; CHAIN: NULL;	CALCIUM-REGULATED MUSCLE CONTRACTION MUSCLE CONTRACTION, CALCIUM-BINDING, TROPONIN, E-F HAND, 2 OPEN CONFORMATION REGULATORY DOMAIN, CALCIUM-REGULATED 3 MUSCLE CONTRACTION
739	1tcf		21	154	1.8e-40	0.15	1.00		TROPONIN C; CHAIN: NULL;	CALCIUM-REGULATED MUSCLE CONTRACTION MUSCLE CONTRACTION, CALCIUM-BINDING, TROPONIN, E-F HAND, 2 OPEN CONFORMATION REGULATORY DOMAIN, CALCIUM-REGULATED 3 MUSCLE CONTRACTION
739	1tnx		16	153	1.8e-36			58.78	TROPONIN C; 1TNX 4 CHAIN: NULL; 1TNX 5	CALCIUM-BINDING PROTEIN EF-HAND 1TNX 14
739	1tnx		21	144	1.8e-36	-0.15	0.88		TROPONIN C; 1TNX 4 CHAIN: NULL; 1TNX 5	CALCIUM-BINDING PROTEIN EF-HAND 1TNX 14
739	1top		21	144	1.8e-40	0.29	1.00		CONTRACTILE SYSTEM PROTEIN TROPONIN C 1TOP 3	
739	1top		3	153	1.8e-40			62.28	CONTRACTILE SYSTEM PROTEIN TROPONIN C 1TOP 3	
739	1vrk	A	18	144	1.6e-51	-0.05	1.00		CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALLING, 2

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	- PDB annotation
										COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
739	1vrk	A	18	153	1.6e-51			66.67	CALMODULIN; CHAIN: A; RS20; CHAIN: B;	CALMODULIN, CALCIUM BINDING, HELIX-LOOP-HELIX, SIGNALLING, 2 COMPLEX(CALCIUM-BINDING PROTEIN/PEPTIDE)
741	1zrn		2	224	1.4e-20	-0.08	0.21		L-2-HALOACID DEHALOGENASE; CHAIN: NULL;	DEHALOGENASE DEHALOGENASE, HYDROLASE
744	1hur	A	2	177	1.8e-62			145.12	HUMAN ADP-RIBOSYLATION FACTOR 1; 1HUR 5 CHAIN: A, B; 1HUR 7	PROTEIN TRANSPORT GDP-BINDING, MEMBRANE TRAFFICKIN, NON-MYRISTOYLATED 1HUR 16
745	1mey	C	678	760	1.6e-48			108.82	DNA; CHAIN: A, B, D, E; CONSENSUS ZINC FINGER PROTEIN; CHAIN: C, F, G;	COMPLEX (ZINC FINGER/DNA) ZINC FINGER, PROTEIN-DNA INTERACTION, PROTEIN DESIGN, 2 CRYSTAL STRUCTURE, COMPLEX (ZINC FINGER/DNA)
745	1tf6	A	408	572	3.6e-37			122.25	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
745	1ubd	C	652	760	1.4e-32			102.06	YY1; CHAIN: C; ADENO-ASSOCIATED VIRUS P5 INITIATOR ELEMENT DNA; CHAIN: A, B;	COMPLEX (TRANSCRIPTION REGULATION/DNA) YING-YANG 1; TRANSCRIPTION INITIATION, INITIATOR ELEMENT, YY1, ZINC 2 FINGER PROTEIN, DNA-PROTEIN RECOGNITION, 3 COMPLEX (TRANSCRIPTION REGULATION/DNA)
753	1hur	A	2	133	7.2e-49			94.14	HUMAN ADP-RIBOSYLATION FACTOR 1; 1HUR 5 CHAIN: A, B; 1HUR 7	PROTEIN TRANSPORT GDP-BINDING, MEMBRANE TRAFFICKIN, NON-MYRISTOYLATED 1HUR 16
757	1hur	A	2	198	1.3e-48			118.23	HUMAN ADP-RIBOSYLATION FACTOR 1; 1HUR 5 CHAIN: A, B;	PROTEIN TRANSPORT GDP-BINDING, MEMBRANE TRAFFICKIN, NON-

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									1HUR 7	MYRISTOYLATED 1HUR 16
762	1klo		618	761	3.6e-25			111.45	LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
767	2occ	L	17	63	7.2e-20			65.51	CYTOCHROME C OXIDASE; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	OXIDOREDUCTASE FERROCYTOCHROME C:OXYGEN OXIDOREDUCTASE; OXIDOREDUCTASE, CYTOCHROME(C)-OXYGEN, CYTOCHROME C 2 OXIDASE
777	2hdc	A	17	113	1.6e-22			126.34	HNF3/FH TRANSCRIPTION FACTOR GENESIS; CHAIN: A; 5'-CHAIN: B; 5'- CHAIN: C;	GENE REGULATION/DNA HEPATOCYTE NUCLEAR FACTOR 3 FORKHEAD HOMOLOG 2, NMR, STRUCTURE, DYANAMICS, GENESIS, WINGED HELIX PROTEIN, 2 GENE REGULATION/DNA
777	2hfh		16	108	1.6e-22			119.17	GENESIS; CHAIN: NULL;	HNF-3 HOMOLOGUES HFH-2; HNF-3 HOMOLOGUES, WINGED HELIX PROTEIN
782	1az9		1	428	0			276.06	AMINOPEPTIDASE P; CHAIN: NULL;	PROLINE PEPTIDASE AMPP; PROLINE PEPTIDASE, HYDROLASE, AMINOPEPTIDASE
782	1c24	A	165	427	1.6e-65			76.10	METHIONINE AMINOPEPTIDASE; CHAIN: A;	HYDROLASE PRODUCT COMPLEX, HYDROLASE
782	1chm	A	3	421	1.1e-57			86.16	CREATINASE CREATINE AMIDINOHYDROLASE (E.C.3.5.3.3) 1CHM 3	
783	1awq	A	1	105	7.2e-56			140.07	CYCLOPHILIN A; CHAIN: A; PEPTIDE FROM THE HIV-1 CAPSID PROTEIN; CHAIN: B;	COMPLEX (ISOMERASE/PEPTIDE) COMPLEX (ISOMERASE/PEPTIDE), CYCLOPHILIN A, HIV-1 CAPSID, 2 PSEUDO-SYMMETRY
786	1a0j	A	192	422	1.8e-94			165.49	TRYPSIN; CHAIN: A, B, C, D;	SERINE PROTEASE SERINE PROTEINASE, TRYPSIN, HYDROLASE
786	1a0l	A	192	423	9e-90			176.56	BETA-TRYPTASE; CHAIN: A, B, C, D;	SERINE PROTEINASE TRYPSIN-LIKE SERINE PROTEINASE, TETRAMER, HEPARIN, ALLERGY, 2 ASTHMA
786	1a5i	A	177	423	3.6e-82			170.05	PLASMINOGEN ACTIVATOR; CHAIN: A; GLU-GLY-ARG	COMPLEX (SERINE PROTEASE/INHIBITOR)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHLOROMETHYL KETONE; CHAIN: I;	(DELTAFEK)DSPAALPHA1; EGRCMK; SERINE PROTEASE, FIBRINOLYTIC ENZYMES, PLASMINOGEN 2 ACTIVATORS
786	1aht	H	192	422	3.6e-77			157.47	ALPHA-THROMBIN; 1AHT 4 CHAIN: L, H; 1AHT 5 HIRUGEN; 1AHT 8 CHAIN: I; 1AHT 9	COMPLEX (SERINE PROTEINASE/INHIBITOR)
786	1aut	C	192	422	7.2e-76			163.21	ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE), PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
786	1bio		192	422	1.3e-69			157.20	COMPLEMENT FACTOR D; CHAIN: NULL;	SERINE PROTEASE SERINE PROTEASE, HYDROLASE, COMPLEMENT, FACTOR D, CATALYTIC 2 TRIAD, SELF-REGULATION
786	1bru	P	192	422	3.6e-90			186.12	ELASTASE; CHAIN: P;	SERINE PROTEASE PPE; SERINE PROTEASE, HYDROLASE
786	1chg		178	423	7.2e-82			166.99	HYDROLASE ZYMOGEN (SERINE PROTEINASE) CHYMOTRYPSINOGEN A 1CHG 4	
786	1dan	H	192	423	3.6e-79			170.73	BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
786	1ekb	B	192	422	5.4e-88			210.66	ENTEROPEPTIDASE; CHAIN: A; ENTEROPEPTIDASE; CHAIN: B; VAL-ASP-ASP-ASP-ASP-LYS PEPTIDE; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR ENTEROKINASE, HEAVY CHAIN; ENTEROKINASE, LIGHT CHAIN; ENTEROPEPTIDASE, TRYPSINOGEN ACTIVATION, 2 HYDROLASE/HYDROLASE INHIBITOR
786	1fxy	A	192	423	5.4e-86			156.85	COAGULATION FACTOR XA-TRYPSIN CHIMERA; CHAIN: A; D-PHE-PRO-ARG-CHLOROMETHYLKETONE (PPACK) WITH CHAIN: I;	COMPLEX (PROTEASE/INHIBITOR) TRYPSIN, COAGULATION FACTOR XA, CHIMERA, PROTEASE, PPACK, 2 CHLOROMETHYLKETONE, COMPLEX (PROTEASE/INHIBITOR)
786	1gct	A	178	423	1.8e-83			174.33	HYDROLASE (SERINE	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									PROTEINASE) GAMMA-*CHYMOTRYPSIN *A (E.C.3.4.21.1) ($P*H 7.0) 1GCT 3	
786	1kig	H	192	423	1.3e-77			168.69	FACTOR XA; CHAIN: H, L; ANTICOAGULANT PEPTIDE; CHAIN: I;	COMPLEX (PROTEASE/INHIBITOR) RTAP; GLYCOPROTEIN, SERINE PROTEASE, PLASMA, BLOOD COAGULATION, 2 COMPLEX (PROTEASE/INHIBITOR)
786	1mct	A	192	423	1.4e-95			159.84	COMPLEX(PROTEINASE/INHIBITOR) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH INHIBITOR FROM BITTER 1MCT 3 GOURD 1MCT 4	
786	1pfx	C	192	423	1.6e-82			177.73	FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
786	1pyt	D	178	423	5.4e-86			170.21	PROCARBOXYPEPTIDASE A; CHAIN: A, B; PROPROTEINASE E; CHAIN: C; CHYMOTRYPSINOGEN C; CHAIN: D;	TERNARY COMPLEX (ZYMOGEN) TC, PCPA-TC; TERNARY COMPLEX (ZYMOGEN), SERINE PROTEINASE, C-TERMINAL 2 PEPTIDASE
786	1qrz	A	176	422	1.8e-95			200.63	PLASMINOGEN; CHAIN: A, B, C, D;	HYDROLASE MICROPLASMINOGEN, SERINE PROTEASE, ZYMOGEN, CHYMOTRYPSIN 2 FAMILY, HYDROLASE
786	1rfn	A	192	423	7.2e-82			174.42	COAGULATION FACTOR IX; CHAIN: A; COAGULATION FACTOR IX; CHAIN: B;	COAGULATION FACTOR SERINE PROTEINASE, BLOOD COAGULATION, COAGULATION FACTOR
786	1rtf	B	192	423	1.8e-81			166.93	TWO CHAIN TISSUE PLASMINOGEN ACTIVATOR; CHAIN: A, B;	SERINE PROTEASE (TC)-T-PA; SERINE PROTEASE, FIBRINOLYTIC ENZYMES
786	1trn	A	192	423	3.6e-93			156.28	HYDROLASE (SERINE PROTEINASE) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH THE INHIBITOR 1TRN 3 DIISOPROPYL-FLUOROPHOSPHOFLUORIDATE	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									(DFP) 1TRN 4 HUMAN TRYPSIN, DFP INHIBITED 1TRN 6	
786	1try		192	421	9e-70			157.71	TRYPSIN; 1TRY 4 CHAIN: NULL; 1TRY 5	HYDROLASE (SERINE PROTEINASE)
786	2tbs		192	423	3.6e-93			160.00	HYDROLASE(SERINE PROTEINASE) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH BENZAMIDINE INHIBITOR 2TBS 3	
786	5ptp		192	423	1.3e-90			153.43	BETA TRYPSIN; CHAIN: NULL;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE, DIGESTION, PANCREAS, 2 ZYMOGEN, SIGNAL
790	1ahd	P	191	258	3.6e-19			77.65	DNA-BINDING PROTEIN ANTENNAPEDIA PROTEIN (HOMEODOMAIN) MUTANT WITH CYS 39 1AHD 3 REPLACED BY SER (C39S) COMPLEX WITH DNA (NMR, 1AHD 4 16 STRUCTURES) 1AHD 5	
790	1b72	A	181	253	5.4e-14			62.50	HOMEOBOX PROTEIN HOX-B1; CHAIN: A; PBX1; CHAIN: B; DNA CHAIN: D; DNA CHAIN: E;	PROTEIN/DNA HOMEODOMAIN, DNA, COMPLEX, DNA-BINDING PROTEIN, PROTEIN/DNA
790	1b8i	A	191	249	1.3e-17			66.24	ULTRABITHORAX HOMEOTIC PROTEIN IV; CHAIN: A; HOMEOBOX PROTEIN EXTRADENTICLE; CHAIN: B; DNA (5'- CHAIN: C; DNA (5'- CHAIN: D;	TRANSCRIPTION/DNA ULTRABITHORAX; PBX PROTEIN; DNA BINDING, HOMEODOMAIN, HOMEOTIC PROTEINS, DEVELOPMENT, 2 SPECIFICITY
790	1ftt		191	258	1.4e-09			58.48	THYROID TRANSCRIPTION FACTOR 1 HOMEODOMAIN; 1FTT 6 CHAIN: NULL; 1FTT 7	DNA BINDING PROTEIN TTF-1 HD; 1FTT 8 DNA BINDING PROTEIN, HOMEODOMAIN, TRANSCRIPTION FACTOR 1FTT 19
790	1ftz		190	259	1.3e-17			70.96	DNA-BINDING FUSHI TARAZU PROTEIN (HOMEODOMAIN) (NMR, 20 STRUCTURES) 1FTZ 3	
790	1san		197	258	1.3e-17			71.71	DNA-BINDING PROTEIN ANTENNAPEDIA PROTEIN (HOMEODOMAIN) MUTANT WITH CYS 39 1SAN 3 REPLACED	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									BY SER AND RESIDUES 1-6 DELETED (C39S,DEL 1-6) 1SAN 4 (NMR, 20 STRUCTURES) 1SAN 5	
790	9ant	A	196	251	1.8e-18			70.57	ANTENNAPEDIA PROTEIN; CHAIN: A, B; DNA; CHAIN: C, D, E, F;	COMPLEX (DNA-BINDING PROTEIN/DNA) HD; HOMEODOMAIN, COMPLEX (DNA-BINDING PROTEIN/DNA)
791	1tf6	A	877	1048	1.1e-37			105.19	TFIIIA; CHAIN: A, D; 5S RIBOSOMAL RNA GENE; CHAIN: B, C, E, F;	COMPLEX (TRANSCRIPTION REGULATION/DNA) COMPLEX (TRANSCRIPTION REGULATION/DNA), RNA POLYMERASE III, 2 TRANSCRIPTION INITIATION, ZINC FINGER PROTEIN
791	2gli	A	844	990	1.6e-35			93.63	ZINC FINGER PROTEIN GLI1; CHAIN: A; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) FIVE-FINGER GLI; GLI, ZINC FINGER, COMPLEX (DNA-BINDING PROTEIN/DNA)
800	1am4	D	18	183	1.6e-42			61.12	P50-RHOGAP; CHAIN: A, B, C; CDC42HS; CHAIN: D, E, F;	COMPLEX (GTPASE-ACTIVATING/GTP-BINDING) COMPLEX (GTPASE-ACTIVATING/GTP-BINDING), GTPASE ACTIVATION
800	1byu	A	16	222	9e-31			63.85	GTP-BINDING PROTEIN RAN; CHAIN: A, B;	TRANSPORT PROTEIN TC4; GTPASE, NUCLEAR TRANSPORT, TRANSPORT PROTEIN
800	1byu	B	12	232	1.3e-31			67.19	GTP-BINDING PROTEIN RAN; CHAIN: A, B;	TRANSPORT PROTEIN TC4; GTPASE, NUCLEAR TRANSPORT, TRANSPORT PROTEIN
800	1cly	A	19	185	5.4e-61			102.61	RAS-RELATED PROTEIN RAP-1A; CHAIN: A; PROTO-ONKOGENE SERINE/THREONINE PROTEIN KINASE CHAIN: B;	SIGNALING PROTEIN GTP-BINDING PROTEINS, PROTEIN-PROTEIN COMPLEX, EFFECTORS
800	1ctq	A	19	186	1.1e-60			101.23	TRANSFORMING PROTEIN P21/H-RAS-1; CHAIN: A;	SIGNALING PROTEIN G PROTEIN, GTP HYDROLYSIS, KINETIC CRYSTALLOGRAPHY, 2 SIGNALING PROTEIN
800	1cxz	A	15	186	1.3e-50			67.32	HIS-TAGGED TRANSFORMING PROTEIN RHOA(0-181); CHAIN: A; PKN; CHAIN: B;	SIGNALING PROTEIN PROTEIN-PROTEIN COMPLEX, ANTIPARALLEL COILED-COIL
800	1ibr	A	19	194	3.6e-30			60.45	RAN; CHAIN: A, C; IMPORTIN	SMALL GTPASE KARYOPHERIN BETA,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									BETA SUBUNIT; CHAIN: B, D;	P95 SMALL GTPASE, NUCLEAR TRANSPORT RECEPTOR
800	1kao		19	186	1.8e-56			113.63	RAP2A; CHAIN: NULL;	GTP-BINDING PROTEIN GTP-BINDING PROTEIN, SMALL G PROTEIN, RAP2, GDP, RAS
800	1mh1		16	191	7.2e-51			72.92	RAC1; CHAIN: NULL;	GTP-BINDING GTP-BINDING, GTPASE, SMALL G-PROTEIN, RHO FAMILY, RAS SUPER 2 FAMILY
800	1rrp	C	18	201	3.6e-30			60.33	RAN; CHAIN: A, C; NUCLEAR PORE COMPLEX PROTEIN NUP358; CHAIN: B, D;	COMPLEX (SMALL GTPASE/NUCLEAR PROTEIN) COMPLEX (SMALL GTPASE/NUCLEAR PROTEIN), SMALL GTPASE, 2 NUCLEAR TRANSPORT
800	1tx4	B	18	183	1.8e-47			56.68	P50-RHOGAP; CHAIN: A; TRANSFORMING PROTEIN RHOA; CHAIN: B;	COMPLEX(GTPASE ACTIVATN/PROTO-ONCOGENE) GTPASE-ACTIVATING PROTEIN RHOGAP; COMPLEX (GTPASE ACTIVATION/PROTO-ONCOGENE), GTPASE, 2 TRANSITION STATE, GAP
800	1zbd	A	17	191	5.4e-55			61.51	RAB-3A; CHAIN: A; RABPHILIN-3A; CHAIN: B;	COMPLEX (GTP-BINDING/EFFECTOR) RAS-RELATED PROTEIN RAB3A; COMPLEX (GTP-BINDING/EFFECTOR), G PROTEIN, EFFECTOR, RABCDR, 2 SYNAPTIC EXOCYTOSIS, RAB PROTEIN, RAB3A, RABPHILIN
800	2ngr	A	19	198	7.2e-46			68.09	GTP BINDING PROTEIN (G25K); CHAIN: A; GTPASE ACTIVATING PROTEIN (RHG); CHAIN: B;	HYDROLASE CDC42/CDC42GAP; CDC42/CDC42GAP; TRANSITION STATE, G-PROTEIN, GAP, CDC42, ALF3., HYDROLASE
800	3rab	A	16	186	1.6e-55			72.05	RAB3A; CHAIN: A;	HYDROLASE G PROTEIN, VESICULAR TRAFFICKING, GTP HYDROLYSIS, RAB 2 PROTEIN, NEUROTRANSMITTER RELEASE, HYDROLASE
814	1bih	A	45	410	7.2e-33			57.74	HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
814	1lil	A	41	247	1.3e-11			52.09	LAMBDA III BENCE JONES PROTEIN CLE; CHAIN: A, B	IMMUNOGLOBULIN IMMUNOGLOBULIN, BENCE JONES PROTEIN
814	1mco	H	1	409	1.8e-39			65.36	IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1)	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									(MCG) WITH A HINGE DELETION IMCO 3	
814	1nfd	F	190	409	7.2e-10			51.22	N15 ALPHA-BETA T-CELL RECEPTOR; CHAIN: A, B, C, D; H57 FAB; CHAIN: E, F, G, H	COMPLEX (IMMUNORECEPTOR/IMMUNOGLOBULIN) COMPLEX (IMMUNORECEPTOR/IMMUNOGLOBULIN)
818	1klo		31	197	1.8e-15			67.17	LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
841	1edh	A	143	350	1.8e-52			91.52	E-CADHERIN; CHAIN: A, B;	CELL ADHESION PROTEIN EPITHELIAL CADHERIN DOMAINS 1 AND 2, ECAD12; CADHERIN, CELL ADHESION PROTEIN, CALCIUM BINDING PROTEIN
843	1a2y	A	37	141	1.8e-35			58.67	MONOCLONAL ANTIBODY D1.3; CHAIN: A, B; LYSOZYME; CHAIN: C;	COMPLEX (IMMUNOGLOBULIN/HYDROLASE) COMPLEX (IMMUNOGLOBULIN/HYDROLASE), IMMUNOGLOBULIN V 2 REGION, SIGNAL, HYDROLASE, GLYCOSIDASE, BACTERIOLYTIC 3 ENZYME, EGG WHITE
843	1a7q	L	37	141	7.2e-33			59.19	MONOCLONAL ANTIBODY D1.3; CHAIN: L, H;	IMMUNOGLOBULIN IMMUNOGLOBULIN, VARIANT
843	1ac6	A	35	143	9e-36			77.29	T-CELL RECEPTOR ALPHA; CHAIN: A, B;	RECEPTOR RECEPTOR, V ALPHA DOMAIN, SITE-DIRECTED MUTAGENESIS, 2 THREE-DIMENSIONAL STRUCTURE, GLYCOPROTEIN, SIGNAL
843	1ao7	D	36	148	9e-39			92.15	HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA-A2 HEAVY CHAIN; CLASS I MHC, T-CELL RECEPTOR, VIRAL PEPTIDE, 2 COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR
843	1ap2	A	37	140	5.4e-35			56.88	MONOCLONAL ANTIBODY C219; CHAIN: A, B, C, D;	IMMUNOGLOBULIN VARIABLE DOMAIN; SINGLE CHAIN FV, MONOCLONAL ANTIBODY, C219, P-GLYCOPROTEIN, 2 IMMUNOGLOBULIN
843	1arl	D	35	141	7.2e-35			66.65	CYTOCHROME C OXIDASE;	COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: A, B; ANTIBODY FV FRAGMENT; CHAIN: C, D;	(OXIDOREDUCTASE/ANTIBODY) CYTOCHROME AA3, COMPLEX IV, FERROCYTOCHROME C, COMPLEX (OXIDOREDUCTASE/ANTIBODY), ELECTRON TRANSPORT, 2 TRANSMEMBRANE, CYTOCHROME OXIDASE, ANTIBODY COMPLEX
843	1b0w	A	35	143	1.1e-37			63.94	BENCE-JONES KAPPA I PROTEIN BRE; CHAIN: A, B, C;	IMMUNE SYSTEM BENCE-JONES; IMMUNOGLOBULIN, AMYLOID, IMMUNE SYSTEM
843	1b88	A	34	143	3.6e-39			75.85	T CELL RECEPTOR V-ALPHA DOMAIN; CHAIN: A, B;	T CELL RECEPTOR TCR; T CELL RECEPTOR, MHC CLASS I, HUMAN IMMUNODEFICIENCY VIRUS, 2 MOLECULAR RECOGNITION
843	1bd2	D	35	167	1.8e-47			60.19	HLA-A 0201; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE; CHAIN: C; T CELL RECEPTOR ALPHA; CHAIN: D; T CELL RECEPTOR BETA; CHAIN: E;	COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR) HLA A2 HEAVY CHAIN; COMPLEX (MHC/VIRAL PEPTIDE/RECEPTOR)
843	1bvk	A	35	143	9e-39			59.82	HULYS11; CHAIN: A, B, D, E; LYSOZYME; CHAIN: C, F;	COMPLEX (HUMANIZED ANTIBODY/HYDROLASE) MURAMIDASE; HUMANIZED ANTIBODY, ANTIBODY COMPLEX, FV, ANTI-LYSOZYME, 2 COMPLEX (HUMANIZED ANTIBODY/HYDROLASE)
843	1bww	A	32	142	9e-40			64.91	IG KAPPA CHAIN V-I REGION REI; CHAIN: A, B;	IMMUNE SYSTEM REIV, STABILIZED IMMUNOGLOBULIN FRAGMENT, BENCE-JONES 2 PROTEIN, IMMUNE SYSTEM
843	1dlf	L	37	143	1.8e-30			56.21	ANTI-DANSYL IMMUNOGLOBULIN IGG2A(S); CHAIN: L, H;	IMMUNOGLOBULIN ANTI-DANSYL FV FRAGMENT FV FRAGMENT, IMMUNOGLOBULIN
843	1fgv	L	35	141	1.8e-40			62.71	IMMUNOGLOBULIN FV FRAGMENT OF A HUMANIZED VERSION OF THE ANTI-CD18 1FGV 3 ANTIBODY 'H52' (HUH52-AA FV) 1FGV 4	
843	1fvc	A	35	144	1.8e-37			59.70	IMMUNOGLOBULIN FV	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									FRAGMENT OF HUMANIZED ANTIBODY 4D5, VERSION 8 1FVC 3	
843	1igm	L	35	149	1.1e-39			57.55	IMMUNOGLOBULIN IMMUNOGLOBULIN M (IG-M) FV FRAGMENT 1IGM 3	
843	1ivl	A	35	141	3.6e-31			62.40	IMMUNOGLOBULIN IMMUNOGLOBULIN VL DOMAIN (VARIABLE DOMAIN OF KAPPA LIGHT 1IVL 3 CHAIN) OF DESIGNED ANTIBODY M29B 1IVL 4	
843	1jhl	L	35	143	3.6e-33			64.05	COMPLEX(ANTIBODY-ANTIGEN) FV FRAGMENT (IGG1, KAPPA) (LIGHT AND HEAVY VARIABLE DOMAINS 1JHL 3 NON-COVALENTLY ASSOCIATED) OF MONOCLONAL ANTI-HEN EGG 1JHL 4 LYSOZYME ANTIBODY D11.15 COMPLEX WITH PHEASANT EGG 1JHL 5 LYSOZYME 1JHL 6	
843	1kb5	A	35	145	3.6e-41			76.59	KB5-C20 T-CELL ANTIGEN RECEPTOR; CHAIN: A, B; ANTIBODY DESIRE-1; CHAIN: L, H;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) TCR VAPLHA VBETA DOMAIN; T-CELL RECEPTOR, STRAND SWITCH, FAB, ANTICLONOTYPIC, 2 (IMMUNOGLOBULIN/RECEPTOR)
843	1qrn	D	36	167	1.8e-44			60.03	MHC CLASS I HLA-A; CHAIN: A; BETA-2 MICROGLOBULIN; CHAIN: B; TAX PEPTIDE P6A; CHAIN: C; HMAN T-CELL RECEPTOR; CHAIN: D; HLA-A 0201; CHAIN: E;	IMMUNE SYSTEM HUMAN TCR/PEPTIDE/MHC COMPLEX, HLA-A2, HTLV-1, TAX, TCR, T 2 CELL RECEPTOR, IMMUNE SYSTEM
843	1rvf	L	36	145	9e-34			62.70	HUMAN RHINOVIRUS 14 COAT PROTEIN; CHAIN: 1, 2, 3, 4; FAB 17-IA; CHAIN: L, H	COMPLEX (COAT PROTEIN/IMMUNOGLOBULIN) POLYPROTEIN, COAT PROTEIN, CORE PROTEIN, RNA-DIRECTED RNA 2

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										POLYMERASE, HYDROLASE, THIOL PROTEASE, MYRISTYLATION, 3 COMPLEX (COAT PROTEIN/IMMUNOGLOBULIN)
843	1tvd	A	35	143	1.1e-20			57.93	T CELL RECEPTOR; CHAIN: A, B;	IMMUNORECEPTOR ES204 V DELTA; IMMUNORECEPTOR, TCR, DELTA CHAIN, VARIABLE DOMAIN
843	1wtl	A	35	143	1.6e-38			59.28	IMMUNOGLOBULIN WAT, A VARIABLE DOMAIN FROM IMMUNOGLOBULIN LIGHT-CHAIN 1WTL 3 (BENCE-JONES PROTEIN) 1WTL 4	
843	2imn		37	143	3.6e-37			58.49	IMUNOGLOBULIN IMMUNOGLOBULIN VL DOMAIN (VARIABLE DOMAIN OF KAPPA 2IMN 3 LIGHT CHAIN) OF MCPC603 MUTANT IN WHICH 2IMN 4 COMPLEMENTARITY-DETERMINING REGION I HAS BEEN REPLACED BY 2IMN 5 THAT FROM MOPC167 2IMN 6	
843	2rhe		35	145	1.8e-41			64.63	IMMUNOGLOBULIN BENCE-*JONES PROTEIN (LAMBDA, VARIABLE DOMAIN) 2RHE 4	
848	1nsy	A	5	130	1.6e-20	0.05	0.07		NAD SYNTHETASE; CHAIN: A, B;	LYASE LYASE, AMIDOTRANSFERASE, NH3 DEPENDENT, ATP PYROPHOSPHATASE
849	1aut	L	32	142	6e-10	0.16	-0.14		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE), PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
849	1c2a	A	4	145	4e-17	0.28	-0.15		BOWMAN-BIRK TRYPSIN INHIBITOR; CHAIN: A	HYDROLASE INHIBITOR ALL-BETA STRUCTURE, HYDROLASE INHIBITOR
849	1ehd	A	31	108	6e-10	0.36	0.66		AGGLUTININ ISOLECTIN VI;	PLANT PROTEIN TWO HOMOLOGOUS

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHAIN: A	HEVEIN-LIKE DOMAINS
849	1ehd	A	81	168	1.4e-07	0.24	0.16		AGGLUTININ ISOLECTIN VI; CHAIN: A	PLANT PROTEIN TWO HOMOLOGOUS HEVEIN-LIKE DOMAINS
849	1eis	A	5	95	2e-08	0.17	0.25		AGGLUTININ ISOLECTIN VI/AGGLUTININ ISOLECTIN V; CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTIGEN
849	1eis	A	76	173	6e-08	0.31	-0.01		AGGLUTININ ISOLECTIN VI/AGGLUTININ ISOLECTIN V; CHAIN: A;	SUGAR BINDING PROTEIN UDA; LECTIN, HEVEIN DOMAIN, UDA, SUPERANTIGEN
849	1ext	A	16	173	1.2e-12			58.35	TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
849	1ext	A	30	171	4e-10	0.21	0.05		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
849	1klo		32	173	1e-17	0.07	-0.09		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
849	1klo		4	140	2e-17	0.12	-0.06		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
849	1klo		4	156	1e-17			62.19	LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
849	1nub	A	79	174	2e-11	0.09	-0.20		BASEMENT MEMBRANE PROTEIN BM-40; CHAIN: A, B;	EXTRACELLULAR MODULE OSTEONECTIN, SPARC, SECRETED PROTEIN ACIDIC AND EXTRACELLULAR MODULE, GLYCOPROTEIN, ANTI-ADHESIVE PROTEIN, 2 COLLAGEN BINDING, SITE-DIRECTED MUTAGENESIS, GLYCOSYLATED 3 PROTEIN MODRES
850	1d5v	A	73	158	1.8e-42	0.32	1.00		S12 TRANSCRIPTION FACTOR (FKH-14); CHAIN: A;	GENE REGULATION WINGED HELIX, DNA-RECOGNITION HELIX
850	1e17	A	69	148	7.2e-27	-0.04	1.00		AFX; CHAIN: A;	DNA BINDING DOMAIN DNA BINDING DOMAIN, WINGED HELIX
850	2hdc	A	73	164	9e-41	0.12	1.00		HNF3/FH TRANSCRIPTION FACTOR GENESIS; CHAIN: A; 5'-CHAIN: B; 5'- CHAIN: C;	GENE REGULATION/DNA HEPATOCYTE NUCLEAR FACTOR 3 FORKHEAD HOMOLOG 2, NMR, STRUCTURE, DYANAMICS, GENESIS, WINGED HELIX PROTEIN, 2 GENE REGULATION/DNA
850	2hdc	A	73	164	9e-41			74.25	HNF3/FH TRANSCRIPTION FACTOR GENESIS; CHAIN: A; 5'-CHAIN: B; 5'- CHAIN: C;	GENE REGULATION/DNA HEPATOCYTE NUCLEAR FACTOR 3 FORKHEAD HOMOLOG 2, NMR, STRUCTURE, DYANAMICS, GENESIS, WINGED HELIX PROTEIN, 2 GENE REGULATION/DNA

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
850	2hfh		73	158	1.6e-39	-0.08	0.94		GENESIS; CHAIN: NULL;	HNF-3 HOMOLOGUES HFH-2; HNF-3 HOMOLOGUES, WINGED HELIX PROTEIN
850	2hfh		73	159	1.6e-39			71.39	GENESIS; CHAIN: NULL;	HNF-3 HOMOLOGUES HFH-2; HNF-3 HOMOLOGUES, WINGED HELIX PROTEIN
855	1a0j	A	561	795	5.4e-98	0.94	1.00		TRYPSIN; CHAIN: A, B, C, D;	SERINE PROTEASE SERINE PROTEINASE, TRYPSIN, HYDROLASE
855	1a0j	A	561	795	5.4e-98			176.29	TRYPSIN; CHAIN: A, B, C, D;	SERINE PROTEASE SERINE PROTEINASE, TRYPSIN, HYDROLASE
855	1a0l	A	561	795	9e-89			206.34	BETA-TRYPTASE; CHAIN: A, B, C, D;	SERINE PROTEINASE TRYPSIN-LIKE SERINE PROTEINASE, TETRAMER, HEPARIN, ALLERGY, 2 ASTHMA
855	1a5i	A	551	795	1.6e-89			191.74	PLASMINOGEN ACTIVATOR; CHAIN: A; GLU-GLY-ARG CHLOROMETHYL KETONE; CHAIN: I;	COMPLEX (SERINE PROTEASE/INHIBITOR) (DELTAFEK)DSPAALPHA1; EGRCMK; SERINE PROTEASE, FIBRINOLYTIC ENZYMES, PLASMINOGEN 2 ACTIVATORS
855	1aht	H	561	795	6e-86			189.27	ALPHA-THROMBIN; 1AHT 4 CHAIN: L, H; 1AHT 5 HIRUGEN; 1AHT 8 CHAIN: I; 1AHT 9	COMPLEX (SERINE PROTEINASE/INHIBITOR)
855	1ajj		441	473	1.6e-09	-0.16	0.54		LOW-DENSITY LIPOPROTEIN RECEPTOR; CHAIN: NULL;	RECEPTOR LR5; RECEPTOR, LDL RECEPTOR, CYSTEINE-RICH MODULE, CALCIUM
855	1aut	C	561	795	1.4e-88			193.08	ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE), PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
855	1aut	L	432	517	3.6e-13	-0.11	0.03		ACTIVATED PROTEIN C; CHAIN: C, L; D-PHE-PRO-MAI; CHAIN: P;	COMPLEX (BLOOD COAGULATION/INHIBITOR) AUTOPROTHROMBIN IIA; HYDROLASE, SERINE PROTEINASE), PLASMA CALCIUM BINDING, 2 GLYCOPROTEIN, COMPLEX (BLOOD COAGULATION/INHIBITOR)
855	1bru	P	561	795	1.1e-90			193.36	ELASTASE; CHAIN: P;	SERINE PROTEASE PPE; SERINE PROTEASE, HYDROLASE
855	1cr8	A	441	473	2e-10	0.30	0.74		LOW DENSITY LIPOPROTEIN	LIPID BINDING PROTEIN RECEPTOR,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									RECEPTOR RELATED PROTEIN; CHAIN: A;	LIGAND BINDING, CALCIUM BINDING, LDLR, LRP, LIPID 2 BINDING PROTEIN
855	1cr8	A	479	509	4e-11	0.17	0.15		LOW DENSITY LIPOPROTEIN RECEPTOR RELATED PROTEIN; CHAIN: A;	LIPID BINDING PROTEIN RECEPTOR, LIGAND BINDING, CALCIUM BINDING, LDLR, LRP, LIPID 2 BINDING PROTEIN
855	1d2j	A	441	474	1.2e-09	-0.43	0.18		LOW-DENSITY LIPOPROTEIN RECEPTOR; CHAIN: A;	SIGNALING PROTEIN LR6*; RECEPTOR, LDLR, CYSTEINE-RICH MODULE, CALCIUM LIGAND- 2 BINDING, FAMILIAL HYPERCHOLESTEROLEMIA
855	1d2l	A	441	476	6e-11	-0.02	0.22		LIPOPROTEIN RECEPTOR RELATED PROTEIN; CHAIN: A;	SIGNALING PROTEIN LIGAND BINDING, CALCIUM BINDING, COMPLEMENT-LIKE REPEAT, 2 RECEPTOR, SIGNALING PROTEIN
855	1d2l	A	514	552	4e-12	0.32	-0.12		LIPOPROTEIN RECEPTOR RELATED PROTEIN; CHAIN: A;	SIGNALING PROTEIN LIGAND BINDING, CALCIUM BINDING, COMPLEMENT-LIKE REPEAT, 2 RECEPTOR, SIGNALING PROTEIN
855	1d6w	A	534	794	1.6e-92	1.00	1.00		THROMBIN; CHAIN: A; DECAPEPTIDE INHIBITOR; CHAIN: I;	HYDROLASE/HYDROLASE INHIBITOR HYDROLASE, THROMBIN, THROMBIN INHIBITOR
855	1dan	H	561	795	2e-81			190.30	BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
855	1dan	L	430	518	1.4e-14	-0.39	0.12		BLOOD COAGULATION FACTOR VIIA; CHAIN: L, H; SOLUBLE TISSUE FACTOR; CHAIN: T, U; D-PHE-PHE-ARG-CHLOROMETHYLKETONE (DFFRCMK) WITH CHAIN: C;	BLOOD COAGULATION, SERINE PROTEASE, COMPLEX, CO-FACTOR, 2 RECEPTOR ENZYME, INHIBITOR, GLA, EGF, 3 COMPLEX (SERINE PROTEASE/COFACTOR/LIGAND)
855	1dva	L	435	518	5.4e-13	-0.21	0.16		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
855	1ekb	B	561	794	8e-92	1.05	1.00		ENTEROPEPTIDASE; CHAIN: A; ENTEROPEPTIDASE; CHAIN: B; VAL-ASP-ASP-ASP-ASP-LYS PEPTIDE; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR ENTEROKINASE, HEAVY CHAIN; ENTEROKINASE, LIGHT CHAIN; ENTEROPEPTIDASE, TRYPSINOGEN ACTIVATION, 2 HYDROLASE/HYDROLASE INHIBITOR
855	1ekb	B	561	795	8e-92			225.15	ENTEROPEPTIDASE; CHAIN: A; ENTEROPEPTIDASE; CHAIN: B; VAL-ASP-ASP-ASP-ASP-LYS PEPTIDE; CHAIN: C;	HYDROLASE/HYDROLASE INHIBITOR ENTEROKINASE, HEAVY CHAIN; ENTEROKINASE, LIGHT CHAIN; ENTEROPEPTIDASE, TRYPSINOGEN ACTIVATION, 2 HYDROLASE/HYDROLASE INHIBITOR
855	1elt		561	794	3.6e-81			180.46	ELASTASE; 1ELT 4 CHAIN: NULL; 1ELT 5	SERINE PROTEINASE
855	1ept	A	561	603	3.6e-17	-0.55	0.99		HYDROLASE (SERINE PROTEASE) PORCINE E-TRYPSIN (E.C.3.4.21.4) 1EPT 3	
855	1ept	A	561	604	8e-19	-0.55	0.90		HYDROLASE (SERINE PROTEASE) PORCINE E-TRYPSIN (E.C.3.4.21.4) 1EPT 3	
855	1etr	H	561	795	4e-86			179.82	HYDROLASE(SERINE PROTEINASE) EPSILON-THROMBIN (E.C.3.4.21.5) NON-COVALENT COMPLEX WITH 1ETR 3 MQPA 1ETR 4	
855	1f5y	A	440	509	1.6e-19	-0.08	0.37		LOW-DENSITY LIPOPROTEIN RECEPTOR; CHAIN: A;	LIPID BINDING PROTEIN LDL RECEPTOR; BETA HAIRPIN, 3-10 HELIX, CALCIUM BINDING
855	1f5y	A	479	550	8e-21	0.41	0.31		LOW-DENSITY LIPOPROTEIN RECEPTOR; CHAIN: A;	LIPID BINDING PROTEIN LDL RECEPTOR; BETA HAIRPIN, 3-10 HELIX, CALCIUM BINDING
855	1f8z	A	443	474	1.2e-09	-0.02	0.57		LOW-DENSITY LIPOPROTEIN RECEPTOR; CHAIN: A;	LIPID BINDING PROTEIN LDL RECEPTOR, LIGAND-BINDING DOMAIN, CALCIUM-BINDING, 2 FAMILIAL HYPERCHOLESTEROLEMIA
855	1fxy	A	561	795	5.4e-89			180.36	COAGULATION FACTOR XA-TRYPSIN CHIMERA; CHAIN: A; D-PHE-PRO-ARG-	COMPLEX (PROTEASE/INHIBITOR) TRYPSIN, COAGULATION FACTOR XA, CHIMERA, PROTEASE, PPACK, 2

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									CHLOROMETHYLKETONE (PPACK) WITH CHAIN: I;	CHLOROMETHYLKETONE, COMPLEX (PROTEASE/INHIBITOR)
855	1gct	A	552	795	5.4e-85			182.37	HYDROLASE (SERINE PROTEINASE) GAMMA-*CHYMOTRYPSIN *A (E.C.3.4.21.1) (SP*H 7.0) 1GCT 3	
855	1kig	H	561	795	1.6e-91			182.52	FACTOR XA; CHAIN: H, L; ANTICOAGULANT PEPTIDE; CHAIN: I;	COMPLEX (PROTEASE/INHIBITOR) RTAP; GLYCOPROTEIN, SERINE PROTEASE, PLASMA, BLOOD COAGULATION, 2 COMPLEX (PROTEASE/INHIBITOR)
855	1ldl		440	476	1.2e-10	-0.01	0.16		LOW-DENSITY LIPOPROTEIN RECEPTOR; 1LDL 4 CHAIN: NULL; 1LDL 5	BINDING PROTEIN LB1; 1LDL 7 LDL RECEPTOR CYSTEINE-RICH REPEAT 1LDL 15
855	1ldl		514	552	2e-12	0.51	0.09		LOW-DENSITY LIPOPROTEIN RECEPTOR; 1LDL 4 CHAIN: NULL; 1LDL 5	BINDING PROTEIN LB1; 1LDL 7 LDL RECEPTOR CYSTEINE-RICH REPEAT 1LDL 15
855	1ldr		441	473	4e-09	-0.16	0.25		LOW-DENSITY LIPOPROTEIN RECEPTOR; 1LDR 5 CHAIN: NULL; 1LDR 6	BINDING PROTEIN LB2; 1LDR 8 LDL RECEPTOR CYSTEINE-RICH REPEAT 1LDR 16
855	1mct	A	561	794	1.8e-99	1.06	1.00		COMPLEX(PROTEINASE/INHIBITOR) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH INHIBITOR FROM BITTER 1MCT 3 GOURD 1MCT 4	
855	1mct	A	561	795	1.8e-99			180.72	COMPLEX(PROTEINASE/INHIBITOR) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH INHIBITOR FROM BITTER 1MCT 3 GOURD 1MCT 4	
855	1mkx	K	523	795	1.8e-89			194.09	ALPHA-THROMBIN; CHAIN: L, H; PRETHROMBIN-2; CHAIN: K;	COMPLEX (BLOOD COAGULATION/PROENZYME) COMPLEX (BLOOD COAGULATION/PROENZYME), THROMBIN, 2 PRETHROMBIN-2, PLASMA, SERINE PROTEASE
855	1pfx	C	561	795	4e-91			186.02	FACTOR IXA; CHAIN: C, L,; D-PHE-PRO-ARG; CHAIN: I;	COMPLEX (BLOOD COAGULATION/INHIBITOR) CHRISTMAS FACTOR; COMPLEX, INHIBITOR, HEMOPHILIA/EGF, BLOOD

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										COAGULATION, 2 PLASMA, SERINE PROTEASE, CALCIUM-BINDING, HYDROLASE, 3 GLYCOPROTEIN
855	1pyt	C	555	795	1.3e-79			177.81	PROCARBOXYPEPTIDASE A; CHAIN: A, B; PROPROTEINASE E; CHAIN: C; CHYMOTRYPSINOGEN C; CHAIN: D;	TERNARY COMPLEX (ZYMOGEN) TC, PCPA-TC; TERNARY COMPLEX (ZYMOGEN), SERINE PROTEINASE, C-TERMINAL 2 PEPTIDASE
855	1pyt	D	550	795	1.8e-82			176.08	PROCARBOXYPEPTIDASE A; CHAIN: A, B; PROPROTEINASE E; CHAIN: C; CHYMOTRYPSINOGEN C; CHAIN: D;	TERNARY COMPLEX (ZYMOGEN) TC, PCPA-TC; TERNARY COMPLEX (ZYMOGEN), SERINE PROTEINASE, C-TERMINAL 2 PEPTIDASE
855	1qrz	A	540	795	5.4e-93			209.13	PLASMINOGEN; CHAIN: A, B, C, D;	HYDROLASE MICROPLASMINOGEN, SERINE PROTEASE, ZYMOGEN, CHYMOTRYPSIN 2 FAMILY, HYDROLASE
855	1qrz	A	550	794	5.4e-93	0.94	1.00		PLASMINOGEN; CHAIN: A, B, C, D;	HYDROLASE MICROPLASMINOGEN, SERINE PROTEASE, ZYMOGEN, CHYMOTRYPSIN 2 FAMILY, HYDROLASE
855	1rfn	A	561	795	1e-90			183.77	COAGULATION FACTOR IX; CHAIN: A; COAGULATION FACTOR IX; CHAIN: B;	COAGULATION FACTOR SERINE PROTEINASE, BLOOD COAGULATION, COAGULATION FACTOR
855	1rtf	B	561	795	1.3e-79			195.82	TWO CHAIN TISSUE PLASMINOGEN ACTIVATOR; CHAIN: A, B;	SERINE PROTEASE (TC)-T-PA; SERINE PROTEASE, FIBRINOLYTIC ENZYMES
855	1slw	B	561	794	3.6e-95	1.06	1.00		ECOTIN; CHAIN: A; ANIONIC TRYPSIN; CHAIN: B;	COMPLEX (SERINE PROTEASE/INHIBITOR) TRYPSIN INHIBITOR; SERINE PROTEASE, INHIBITOR, COMPLEX, METAL BINDING SITES, 2 PROTEIN ENGINEERING, PROTEASE-SUBSTRATE INTERACTIONS, 3 METALLOPROTEINS
855	1trn	A	561	794	1.3e-96	1.04	1.00		HYDROLASE (SERINE PROTEINASE) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH THE INHIBITOR 1TRN 3 DIISOPROPYL-FLUOROPHOSPHOFLUORIDATE (DFP) 1TRN 4 HUMAN TRYPSIN, DFP INHIBITED 1TRN 6	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
855	1uvu	H	561	794	1.1e-75			177.88	THROMBIN; CHAIN: L, H;	SERINE PROTEASE FACTOR II; SERINE PROTEASE, HYDROLASE, THROMBIN, BLOOD COAGULATION
855	2tbs		561	794	1.6e-96	1.00	1.00		HYDROLASE(SERINE PROTEINASE) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH BENZAMIDINE INHIBITOR 2TBS 3	
855	5ptp		561	794	1.8e-94	1.01	1.00		BETA TRYPSIN; CHAIN: NULL;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE, DIGESTION, PANCREAS, 2 ZYMOGEN, SIGNAL
855	5ptp		561	795	1.8e-94			176.17	BETA TRYPSIN; CHAIN: NULL;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE, DIGESTION, PANCREAS, 2 ZYMOGEN, SIGNAL
858	1ccd		30	91	0.00012	0.17	0.82		PHOSPHOLIPASE A2 INHIBITOR CLARA CELL 17-KDA PROTEIN 1CCD 3	
858	1ccd		30	91	3.6e-12	-0.28	0.53		PHOSPHOLIPASE A2 INHIBITOR CLARA CELL 17-KDA PROTEIN 1CCD 3	
858	1utg		30	91	1.8e-05	0.01	0.45		STEROID BINDING UTEROGLOBIN (OXIDIZED) 1UTG 4	
858	1utr	A	30	91	0.00012	0.03	0.63		UTEROGLOBIN; 1UTR 5 CHAIN: A, B; 1UTR 6	MAMMALIAN PCB-BINDING PROTEIN MAMMALIAN PCB-BINDING PROTEIN;. - 1UTR 7 UTEROGLOBIN, CLARA CELL 17 KDA PROTEIN (CC10), 1UTR 18 2 PHOSPHOLIPASE A2 INHIBITOR, CLARA CELL PHOSPHOLIPID-BINDING 1UTR 19 3 PROTEIN, PROGESTERONE BINDING 1UTR 20
860	1cwn		2	325	0	0.24	1.00		ALDEHYDE REDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE ALR1; TIM-BARREL, OXIDOREDUCTASE, NADP
860	1cwn		2	325	0			537.02	ALDEHYDE REDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE ALR1; TIM-BARREL, OXIDOREDUCTASE, NADP
860	2alr		2	325	0	0.87	1.00		ALDEHYDE REDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE ALR1; OXIDOREDUCTASE, TIM-BARREL

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
860	2alr		2	325	0			505.09	ALDEHYDE REDUCTASE; CHAIN: NULL;	OXIDOREDUCTASE ALR1; OXIDOREDUCTASE, TIM-BARREL
861	1914		14	115	0.0027	-0.30	0.07		SIGNAL RECOGNITION PARTICLE 9/14 FUSION PROTEIN; CHAIN: NULL;	ALU DOMAIN SRP9/14, ALU BM, RBD; ALU DOMAIN, CRYSTAL STRUCTURE, RNA BINDING, SIGNAL 2 RECOGNITION PARTICLE (SRP), TRANSLATION REGULATION
861	1e8o	B	14	49	0.0072	-0.63	0.33		SIGNAL RECOGNITION PARTICLE 9 KDA PROTEIN; CHAIN: A, C; SIGNAL RECOGNITION PARTICLE 14 KDA PROTEIN; CHAIN: B, D; 7SL RNA, 5'-R(GDP*GP*GP*CP*CP*GP*GP*GP*CP*GP*CP*GP* CHAIN: E;	ALU RIBONUCLEOPROTEIN PARTICLE SRP9; SRP14; ALU RIBONUCLEOPROTEIN PARTICLE, PROTEIN RECOGNITION OF AN 2 RNA U-TURN, TRANSLATIONAL CONTROL, ALU RNP ASSEMBLY AND 3 TRANSPORT, ALU RETROPOSITION
863	1du3	B	379	467	1.1e-08	0.14	0.06		DEATH RECEPTOR 5; CHAIN: A, B, C, G, H, I; TNF-RELATED APOPTOSIS INDUCING LIGAND; CHAIN: D, E, F, J, K, L;	APOPTOSIS TRAIL, DR5, COMPLEX
863	1ext	A	121	258	1.7e-11	0.14	-0.15		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
863	1ext	A	64	192	1.8e-08	-0.07	0.06		TUMOR NECROSIS FACTOR RECEPTOR; CHAIN: A, B;	SIGNALLING PROTEIN BINDING PROTEIN, CYTOKINE, SIGNALLING PROTEIN
863	1ezg	A	132	212	3.4e-07	0.16	0.06		THERMAL HYSTERESIS PROTEIN ISOFORM YL-1; CHAIN: A, B;	ANTIFREEZE PROTEIN INSECT ANTIFREEZE PROTEIN, THERMAL HYSTERESIS, TENEBRIO 2 MOLITOR, IODINATION, RIGHT-HANDED BETA-HELIX, TMAFP
863	1igr	A	105	253	3.4e-08	0.18	-0.19		INSULIN-LIKE GROWTH FACTOR RECEPTOR 1; CHAIN: A;	HORMONE RECEPTOR HORMONE RECEPTOR, INSULIN RECEPTOR FAMILY
863	1klo		348	496	5.1e-11	0.07	-0.19		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
863	1klo		45	206	5.1e-12	0.03	-0.19		LAMININ; CHAIN: NULL;	GLYCOPROTEIN GLYCOPROTEIN
863	1ncf	A	365	466	3.6e-07	0.23	0.41		TUMOR NECROSIS FACTOR RECEPTOR; 1NCF 4 CHAIN: A, B; 1NCF 5	SIGNALLING PROTEIN TYPE I RECEPTOR, STNFR1; 1NCF 8 BINDING PROTEIN, CYTOKINE 1NCF 19
863	1ncf	A	64	190	1.8e-10	0.17	-0.15		TUMOR NECROSIS FACTOR	SIGNALLING PROTEIN TYPE I RECEPTOR,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									RECEPTOR; 1NCF 4 CHAIN: A, B; 1NCF 5	STNFR1; 1NCF 8 BINDING PROTEIN, CYTOKINE 1NCF 19
863	4mt2		142	209	3.4e-09	0.27	-0.19		METALLOTHIONEIN METALLOTHIONEIN ISOFORM II 4MT2 3	
863	4mt2		532	586	5.1e-08	0.20	-0.13		METALLOTHIONEIN METALLOTHIONEIN ISOFORM II 4MT2 3	
863	9wga	A	102	273	5.1e-16	0.02	-0.19		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
864	1a0j	A	318	538	3.4e-46			103.88	TRYPSIN; CHAIN: A, B, C, D;	SERINE PROTEASE SERINE PROTEINASE, TRYPSIN, HYDROLASE
864	1a0j	A	394	538	3.4e-46	-0.01	0.98		TRYPSIN; CHAIN: A, B, C, D;	SERINE PROTEASE SERINE PROTEINASE, TRYPSIN, HYDROLASE
864	1aks	B	436	538	1.7e-43	0.22	1.00		ALPHA TRYPSIN; CHAIN: A, B;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE
864	1bru	P	322	538	5.1e-39			90.52	ELASTASE; CHAIN: P;	SERINE PROTEASE PPE; SERINE PROTEASE, HYDROLASE
864	1chg		306	537	8.5e-35			103.24	HYDROLASE ZYMOGEN (SERINE PROTEINASE) CHYMOTRYPSINOGEN A 1CHG 4	
864	1ejn	A	422	535	1.8e-43	0.27	0.98		UROKINASE-TYPE PLASMINOGEN ACTIVATOR; CHAIN: A;	HYDROLASE HUMAN, UPA, PLASMINOGEN ACTIVATOR, UROKINASE, INHIBITOR 2 COMPLEX
864	1fiz	A	434	540	5.4e-42	0.09	0.95		BETA-ACROSIN HEAVY CHAIN; CHAIN: A; BETA-ACROSIN LIGHT CHAIN; CHAIN: L	HYDROLASE ANTI-PARALLEL BETA-BARREL
864	1fxy	A	318	539	1.7e-45			102.23	COAGULATION FACTOR XA-TRYPSIN CHIMERA; CHAIN: A; D-PHE-PRO-ARG-CHLOROMETHYLKETONE (PPACK) WITH CHAIN: I;	COMPLEX (PROTEASE/INHIBITOR) TRYPSIN, COAGULATION FACTOR XA, CHIMERA, PROTEASE, PPACK, 2 CHLOROMETHYLKETONE, COMPLEX (PROTEASE/INHIBITOR)
864	1fxy	A	345	538	1.7e-45	0.02	0.55		COAGULATION FACTOR XA-TRYPSIN CHIMERA; CHAIN: A; D-PHE-PRO-ARG-CHLOROMETHYLKETONE	COMPLEX (PROTEASE/INHIBITOR) TRYPSIN, COAGULATION FACTOR XA, CHIMERA, PROTEASE, PPACK, 2 CHLOROMETHYLKETONE, COMPLEX

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									(PPACK) WITH CHAIN: I;	(PROTEASE/INHIBITOR)
864	1gct	A	306	538	1.7e-34			106.80	HYDROLASE (SERINE PROTEINASE) GAMMA-*CHYMOTRYPSIN *A (E.C.3.4.21.1) ($P*H 7.0) 1GCT 3	
864	1mct	A	318	538	3.4e-46			97.39	COMPLEX(PROTEINASE/INHIBITOR) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH INHIBITOR FROM BITTER 1MCT 3 GOURD 1MCT 4	
864	1mct	A	381	538	3.4e-46	0.26	0.99		COMPLEX(PROTEINASE/INHIBITOR) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH INHIBITOR FROM BITTER 1MCT 3 GOURD 1MCT 4	
864	1qrz	A	436	538	1.3e-42	0.11	1.00		PLASMINOGEN; CHAIN: A, B, C, D;	HYDROLASE MICROPLASMINOGEN, SERINE PROTEASE, ZYMOGEN, CHYMOTRYPSIN 2 FAMILY, HYDROLASE
864	1sgf	G	296	539	1.4e-39			96.16	NERVE GROWTH FACTOR; CHAIN: A, B, G, X, Y, Z;	GROWTH FACTOR 7S NGF; GROWTH FACTOR (BETA-NGF), HYDROLASE - SERINE PROTEINASE 2 (GAMMA-NGF), INACTIVE SERINE PROTEINASE (ALPHA-NGF)
864	1slw	B	331	538	1.7e-44			96.11	ECOTIN; CHAIN: A; ANIONIC TRYPSIN; CHAIN: B;	COMPLEX (SERINE PROTEASE/INHIBITOR) TRYPSIN INHIBITOR; SERINE PROTEASE, INHIBITOR, COMPLEX, METAL BINDING SITES, 2 PROTEIN ENGINEERING, PROTEASE-SUBSTRATE INTERACTIONS, 3 METALLOPROTEINS
864	1slw	B	394	538	1.7e-44	0.22	0.80		ECOTIN; CHAIN: A; ANIONIC TRYPSIN; CHAIN: B;	COMPLEX (SERINE PROTEASE/INHIBITOR) TRYPSIN INHIBITOR; SERINE PROTEASE, INHIBITOR, COMPLEX, METAL BINDING SITES, 2 PROTEIN ENGINEERING, PROTEASE-SUBSTRATE INTERACTIONS, 3 METALLOPROTEINS
864	1ton		303	539	1.5e-34			92.41	HYDROLASE(SERINE	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									PROTEINASE) TONIN (E.C. NUMBER NOT ASSIGNED) 1TON 4	
864	1trn	A	318	539	3.4e-45			101.52	HYDROLASE (SERINE PROTEINASE) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH THE INHIBITOR 1TRN 3 DIISOPROPYL-FLUOROPHOSPHOFLUORIDATE (DFP) 1TRN 4 HUMAN TRYPSIN, DFP INHIBITED 1TRN 6	
864	1trn	A	394	538	3.4e-45	0.16	0.93		HYDROLASE (SERINE PROTEINASE) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH THE INHIBITOR 1TRN 3 DIISOPROPYL-FLUOROPHOSPHOFLUORIDATE (DFP) 1TRN 4 HUMAN TRYPSIN, DFP INHIBITED 1TRN 6	
864	2tbs		300	538	3.4e-41			98.98	HYDROLASE(SERINE PROTEINASE) TRYPSIN (E.C.3.4.21.4) COMPLEXED WITH BENZAMIDINE INHIBITOR 2TBS 3	
864	5ptp		318	538	1.7e-43			100.26	BETA TRYPSIN; CHAIN: NULL;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE, DIGESTION, PANCREAS, 2 ZYMOGEN, SIGNAL
864	5ptp		394	538	1.7e-43	0.16	1.00		BETA TRYPSIN; CHAIN: NULL;	SERINE PROTEASE HYDROLASE, SERINE PROTEASE, DIGESTION, PANCREAS, 2 ZYMOGEN, SIGNAL
867	1agi		205	302	0.0054	0.21	0.28		ANGIOGENIN; 1AGI 4 CHAIN: NULL; 1AGI 5	ENDONUCLEASE
867	1b1i	A	205	302	0.00036	0.14	0.21		HYDROLASE ANGIOGENIN; CHAIN: A;	HYDROLASE HYDROLASE (VASCULARIZATION)
867	1rnf	A	205	308	0.0054	0.26	0.68		RIBONUCLEASE 4; CHAIN: A, B;	HYDROLASE RNASE 4; HYDROLASE, RIBONUCLEASE, PHOSPHODIESTERASE
868	1d0s	A	232	430	1.4e-21	0.05	-0.17		NICOTINATE MONONUCLEOTIDE:5,6- CHAIN:	TRANSFERASE DINUCLEOTIDE-BINDING MOTIF, PHOSPHORIBOSYL TRANSFERASE

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									A;	
869	1b72	A	99	156	5.1e-28	-0.22	0.36		HOMEOBOX PROTEIN HOX-B1; CHAIN: A; PBX1; CHAIN: B; DNA CHAIN: D; DNA CHAIN: E;	PROTEIN/DNA HOMEODOMAIN, DNA, COMPLEX, DNA-BINDING PROTEIN, PROTEIN/DNA
869	1b8i	A	99	153	5.1e-28	-0.24	0.24		ULTRABITHORAX HOMEOTIC PROTEIN IV; CHAIN: A; HOMEOBOX PROTEIN EXTRADENTICLE; CHAIN: B; DNA (5'- CHAIN: C; DNA (5'- CHAIN: D;	TRANSCRIPTION/DNA ULTRABITHORAX; PBX PROTEIN; DNA BINDING, HOMEODOMAIN, HOMEOTIC PROTEINS, DEVELOPMENT, 2 SPECIFICITY
869	1du0	B	98	152	1.2e-21	0.06	0.48		ENGRAILED HOMEODOMAIN; CHAIN: A, B; DNA (5'- CHAIN: C; DNA (5'- CHAIN: D;	TRANSCRIPTION/DNA HOMEOTIC PROTEIN ENGRAILED, SEGMENTATION POLARITY HOMEODOMAIN, DNA-BINDING PROTEIN, PROTEIN-DNA COMPLEX
869	1enh		98	149	3.4e-21	-0.00	0.90		DNA-BINDING PROTEIN ENGRAILED HOMEODOMAIN 1ENH 3	
869	1ftz		98	153	1e-27	-0.06	0.17		DNA-BINDING FUSHI TARAZU PROTEIN (HOMEODOMAIN) (NMR, 20 STRUCTURES) 1FTZ 3	
869	1san		101	155	8.5e-31	-0.31	0.04		DNA-BINDING PROTEIN ANTENNAPEDIA PROTEIN (HOMEODOMAIN) MUTANT WITH CYS 39 1SAN 3 REPLACED BY SER AND RESIDUES 1-6 DELETED (C39S,DEL 1-6) 1SAN 4 (NMR, 20 STRUCTURES) 1SAN 5	
869	2hdd	A	99	152	1.7e-21	0.13	0.80		ENGRAILED HOMEODOMAIN; CHAIN: A, B; DNA (20-MER); CHAIN: C, D;	COMPLEX (DNA BINDING PROTEIN/DNA) DNA BINDING, COMPLEX (DNA BINDING PROTEIN/DNA)
869	2hdd	B	98	151	3.4e-21	-0.10	0.58		ENGRAILED HOMEODOMAIN; CHAIN: A, B; DNA (20-MER); CHAIN: C, D;	COMPLEX (DNA BINDING PROTEIN/DNA) DNA BINDING, COMPLEX (DNA BINDING PROTEIN/DNA)
869	9ant	A	99	154	5.1e-31	-0.18	0.11		ANTENNAPEDIA PROTEIN; CHAIN: A, B; DNA; CHAIN: C, D, E, F;	COMPLEX (DNA-BINDING PROTEIN/DNA) HD; HOMEODOMAIN, COMPLEX (DNA-BINDING PROTEIN/DNA)

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
876	1a7i		34	92	6.8e-14	-0.26	0.17		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
876	1a7i		36	92	1.8e-20	-0.24	0.58		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
876	1a7i		94	151	7.2e-12	0.06	1.00		QCRP2 (LIM1); CHAIN: NULL;	LIM DOMAIN CONTAINING PROTEINS LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN, ZINC 2 FINGER
876	1au7	A	146	221	9e-22	0.36	0.98		PIT-1; CHAIN: A, B; DNA; CHAIN: C, D;	COMPLEX (DNA-BINDING PROTEIN/DNA) GHF-1; COMPLEX (DNA-BINDING PROTEIN/DNA), PITUITARY, CPHD, 2 POU DOMAIN, TRANSCRIPTION FACTOR
876	1b72	A	163	223	9e-11	0.38	0.99		HOMEOBOX PROTEIN HOX-B1; CHAIN: A; PBX1; CHAIN: B; DNA CHAIN: D; DNA CHAIN: E;	PROTEIN/DNA HOMEODOMAIN, DNA, COMPLEX, DNA-BINDING PROTEIN, PROTEIN/DNA
876	1bw5		160	225	5.4e-21			50.75	INSULIN GENE ENHANCER PROTEIN ISL-1; CHAIN: NULL;	DNA-BINDING PROTEIN ISL-1HD DNA-BINDING PROTEIN, HOMEODOMAIN, LIM DOMAIN
876	1bw5		163	222	5.4e-21	0.11	0.84		INSULIN GENE ENHANCER PROTEIN ISL-1; CHAIN: NULL;	DNA-BINDING PROTEIN ISL-1HD DNA-BINDING PROTEIN, HOMEODOMAIN, LIM DOMAIN
876	1ctl		34	87	3.6e-18	0.13	0.54		AVIAN CYSTEINE RICH PROTEIN; 1CTL 3	METAL-BINDING PROTEIN LIM DOMAIN CONTAINING PROTEINS 1CTL 15
876	1ctl		36	101	3.4e-14	-0.43	0.07		AVIAN CYSTEINE RICH PROTEIN; 1CTL 3	METAL-BINDING PROTEIN LIM DOMAIN CONTAINING PROTEINS 1CTL 15
876	1ctl		90	150	1.8e-12	-0.22	0.83		AVIAN CYSTEINE RICH PROTEIN; 1CTL 3	METAL-BINDING PROTEIN LIM DOMAIN CONTAINING PROTEINS 1CTL 15
876	1cxx	A	34	89	5.4e-18	0.11	0.41		CYSTEINE AND GLYCINE-RICH PROTEIN CRP2; CHAIN: A;	SIGNALING PROTEIN LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN
876	1cxx	A	35	89	1.7e-11	0.14	0.65		CYSTEINE AND GLYCINE-RICH PROTEIN CRP2; CHAIN: A;	SIGNALING PROTEIN LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
876	1cxx	A	94	151	1.1e-11	-0.40	1.00		CYSTEINE AND GLYCINE-RICH PROTEIN CRP2; CHAIN: A;	SIGNALING PROTEIN LIM DOMAIN CONTAINING PROTEINS, METAL-BINDING PROTEIN
876	1du0	B	164	219	3.6e-11	0.86	1.00		ENGRAILED HOMEODOMAIN; CHAIN: A, B; DNA (5'- CHAIN: C; DNA (5'- CHAIN: D;	TRANSCRIPTION/DNA HOMEOTIC PROTEIN ENGRAILED, SEGMENTATION POLARITY HOMEODOMAIN, DNA-BINDING PROTEIN, PROTEIN-DNA COMPLEX
876	1fjl	A	163	222	5.4e-15	0.60	1.00		PAIRED PROTEIN; CHAIN: A, B, C; DNA; CHAIN: D, E, F	COMPLEX (DNA-BINDING PROTEIN/DNA) DNA-BINDING PROTEIN, DNA, PAIRED BOX, TRANSCRIPTION 2 REGULATION
876	1fjl	B	163	219	1.8e-15	0.55	1.00		PAIRED PROTEIN; CHAIN: A, B, C; DNA; CHAIN: D, E, F	COMPLEX (DNA-BINDING PROTEIN/DNA) DNA-BINDING PROTEIN, DNA, PAIRED BOX, TRANSCRIPTION 2 REGULATION
876	1ftt		163	227	9e-11	0.41	0.82		THYROID TRANSCRIPTION FACTOR 1 HOMEODOMAIN; 1FTT 6 CHAIN: NULL; 1FTT 7	DNA BINDING PROTEIN TTF-1 HD; 1FTT 8 DNA BINDING PROTEIN, HOMEODOMAIN, TRANSCRIPTION FACTOR 1FTT 19
876	1iml		34	105	3.6e-26	0.02	0.54		CYSTEINE RICH INTESTINAL PROTEIN; CHAIN: NULL;	METAL-BINDING PROTEIN CRIP; METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
876	1iml		34	94	5.1e-14	0.33	0.81		CYSTEINE RICH INTESTINAL PROTEIN; CHAIN: NULL;	METAL-BINDING PROTEIN CRIP; METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
876	1iml		94	150	5.4e-11	-0.18	0.95		CYSTEINE RICH INTESTINAL PROTEIN; CHAIN: NULL;	METAL-BINDING PROTEIN CRIP; METAL-BINDING PROTEIN, LIM DOMAIN PROTEIN
876	1nk2	P	163	225	9e-11	0.57	0.95		HOMEOBOX PROTEIN VND; CHAIN: P; DNA; CHAIN: A, B;	COMPLEX (HOMEODOMAIN/DNA) VND/NK-2 HOMEODOMAIN, VENTRAL NERVOUS SYSTEM HOMEODOMAIN, HOMEOBOX, DNA-BINDING PROTEIN, EMBRYONIC 2 DEVELOPMENT, COMPLEX (HOMEODOMAIN/DNA)
876	1ocp		163	219	1.3e-21	0.56	0.99		OCT-3; 1OCP 5 CHAIN: NULL; 1OCP 6	DNA-BINDING PROTEIN
876	1zfo		34	61	6.8e-06	-0.15	0.43		LASP-1; CHAIN: NULL;	METAL-BINDING PROTEIN LIM DOMAIN, ZINC-FINGER, METAL-BINDING PROTEIN
882	1crz	A	1075	1283	0.0018	0.07	0.52		TOLB PROTEIN; CHAIN: A;	TOXIN BINDING PROTEIN TWO

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										DOMAINS: BETA PROPELLER AND ALPHA/BETA FOLD
882	1crz	A	1206	1420	3.4e-06	0.58	0.07		TOLB PROTEIN; CHAIN: A;	TOXIN BINDING PROTEIN TWO DOMAINS: BETA PROPELLER AND ALPHA/BETA FOLD
882	1erj	A	1012	1315	3.4e-60	0.39	0.55		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
882	1erj	A	1092	1433	6.8e-57	0.49	0.82		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
882	1erj	A	1186	1483	3.4e-57	0.10	-0.11		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
882	1erj	A	1234	1574	8.5e-59	0.09	-0.19		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
882	1erj	A	957	1277	5.1e-54	0.09	0.76		TRANSCRIPTIONAL REPRESSOR TUP1; CHAIN: A, B, C;	TRANSCRIPTION INHIBITOR BETA-PROPELLER
882	1got	B	1010	1314	1.2e-71	0.60	1.00		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
882	1got	B	949	1274	6.8e-52	0.39	0.27		GT-ALPHA/GI-ALPHA CHIMERA; CHAIN: A; GT-BETA; CHAIN: B; GT-GAMMA; CHAIN: G;	COMPLEX (GTP-BINDING/TRANSDUCER) BETA1, TRANSDUCIN BETA SUBUNIT; GAMMA1, TRANSDUCIN GAMMA SUBUNIT; COMPLEX (GTP-BINDING/TRANSDUCER), G PROTEIN, HETEROTRIMER 2 SIGNAL TRANSDUCTION
882	2mta	H	1178	1250	0.0058	-0.07	0.25		ELECTRON TRANSPORT METHYLAMINE DEHYDROGENASE (E.C.1.4.99.3) COMPLEX WITH 2MTA 3 AMICYANIN AND CYTOCHROME C551I 2MTA 4	
883	1d8d	A	23	400	1.7e-35	-0.61	0.03		FARNESYLTRANSFERASE (ALPHA SUBUNIT); CHAIN: A;	TRANSFERASE FTASE; FTASE; FTASE, PFT, PFTASE, FARNESYLTRANSFERASE,

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									FARNESYLTRANSFERASE (BETA SUBUNIT); CHAIN: B; K-RAS4B PEPTIDE SUBSTRATE; CHAIN: P;	FARNESYL 2 TRANSFERASE, CAAX, RAS, CANCER
883	1dce	A	57	327	1e-20	-0.45	0.37		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
884	1d8d	A	23	400	1.7e-35	-0.61	0.03		FARNESYLTRANSFERASE (ALPHA SUBUNIT); CHAIN: A; FARNESYLTRANSFERASE (BETA SUBUNIT); CHAIN: B; K-RAS4B PEPTIDE SUBSTRATE; CHAIN: P;	TRANSFERASE FTASE; FTASE; FTASE, PFT, PFTASE, FARNESYLTRANSFERASE, FARNESYL 2 TRANSFERASE, CAAX, RAS, CANCER
884	1dce	A	57	327	1e-20	-0.45	0.37		RAB GERANYLGERANYLTRANSFERASE ALPHA SUBUNIT; CHAIN: A, C; RAB GERANYLGERANYLTRANSFERASE BETA SUBUNIT; CHAIN: B, D;	TRANSFERASE CRYSTAL STRUCTURE, RAB GERANYLGERANYLTRANSFERASE, 2.0 A 2 RESOLUTION, N-FORMYLMETHIONINE, ALPHA SUBUNIT, BETA SUBUNIT
886	2occ	L	17	63	5.1e-21			61.47	CYTOCHROME C OXIDASE; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	OXIDOREDUCTASE FERROCYTOCHROME C\:OXYGEN OXIDOREDUCTASE; OXIDOREDUCTASE, CYTOCHROME(C)-OXYGEN, CYTOCHROME C 2 OXIDASE
886	2occ	L	18	62	5.1e-21	-0.47	0.90		CYTOCHROME C OXIDASE; CHAIN: A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q,	OXIDOREDUCTASE FERROCYTOCHROME C\:OXYGEN OXIDOREDUCTASE; OXIDOREDUCTASE, CYTOCHROME(C)-OXYGEN, CYTOCHROME C 2 OXIDASE
889	1bih	A	40	446	3.4e-41			77.07	HEMOLIN; CHAIN: A, B;	INSECT IMMUNITY INSECT IMMUNITY, LPS-BINDING, HOMOPHILIC ADHESION
889	1cvs	C	144	348	3.4e-47	-0.29	0.34		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
										RECEPTOR
889	1cvs	D	144	348	3.4e-44	-0.39	0.37		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
889	1cvs	D	30	231	1.4e-30	-0.21	0.06		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF, FGFR, IMMUNOGLOBULIN-LIKE, SIGNAL TRANSDUCTION, 2 DIMERIZATION, GROWTH FACTOR/GROWTH FACTOR RECEPTOR
889	1ev2	E	145	348	5.1e-39	-0.54	0.03		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG)LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
889	1ev2	G	145	352	5.1e-42	-0.38	0.12		FIBROBLAST GROWTH FACTOR 2; CHAIN: A, B, C, D; FIBROBLAST GROWTH FACTOR RECEPTOR 2; CHAIN: E, F, G, H;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF2; FGFR2; IMMUNOGLOBULIN (IG)LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
889	1evt	C	141	348	6.8e-43	-0.30	0.10		FIBROBLAST GROWTH FACTOR 1; CHAIN: A, B; FIBROBLAST GROWTH FACTOR RECEPTOR 1; CHAIN: C, D;	GROWTH FACTOR/GROWTH FACTOR RECEPTOR FGF1; FGFR1; IMMUNOGLOBULIN (IG) LIKE DOMAINS BELONGING TO THE I-SET 2 SUBGROUP WITHIN IG-LIKE DOMAINS, B-TREFOIL FOLD
889	1fhg	A	237	348	6.8e-15	0.22	0.69		TELOKIN; CHAIN: A	CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA BARREL
889	1fhg	A	30	133	5.1e-16	0.34	0.10		TELOKIN; CHAIN: A	CONTRACTILE PROTEIN IMMUNOGLOBULIN FOLD, BETA BARREL
889	1fyv	A	389	562	1.8e-34	0.55	0.98		TOLL-LIKE RECEPTOR 1; CHAIN: A;	SIGNALING PROTEIN BETA-ALPHA-BETA FOLD PARALLEL BETA SHEET
889	1fyx	A	400	557	1.8e-26	0.01	0.80		TOLL-LIKE RECEPTOR 2; CHAIN:	SIGNALING PROTEIN BETA-ALPHA-BETA

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									A;	FOLD
889	1igy	B	36	446	3.4e-30			66.45	IGG1 INTACT ANTIBODY MAB61.1.3; CHAIN: A, B, C, D	IMMUNOGLOBULIN INTACT IMMUNOGLOBULIN, V REGION, C REGION, HINGE REGION
889	1itb	B	41	356	3.6e-47			164.52	INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
889	1itb	B	46	346	3.6e-47	-0.13	1.00		INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
889	1itb	B	73	350	5.1e-40	-0.18	1.00		INTERLEUKIN-1 BETA; CHAIN: A; TYPE 1 INTERLEUKIN-1 RECEPTOR; CHAIN: B;	COMPLEX (IMMUNOGLOBULIN/RECEPTOR) IMMUNOGLOBULIN FOLD, TRANSMEMBRANE, GLYCOPROTEIN, RECEPTOR, 2 SIGNAL, COMPLEX (IMMUNOGLOBULIN/RECEPTOR)
889	1mco	H	22	446	1e-32			80.65	IMMUNOGLOBULIN IMMUNOGLOBULIN G1 (IGG1) (MCG) WITH A HINGE DELETION 1MCO 3	
889	1nct		147	232	1.7e-16	-0.13	0.12		TITIN; CHAIN: NULL;	MUSCLE PROTEIN CONNECTIN, NEXTM5; CELL ADHESION, GLYCOPROTEIN, TRANSMEMBRANE, REPEAT, BRAIN, 2 IMMUNOGLOBULIN FOLD, ALTERNATIVE SPLICING, SIGNAL, 3 MUSCLE PROTEIN
889	1nfd	E	149	343	3.4e-15	-0.30	0.05		N15 ALPHA-BETA T-CELL RECEPTOR; CHAIN: A, B, C, D; H57 FAB; CHAIN: E, F, G, H	COMPLEX (IMMUNORECEPTOR/IMMUNOGLOBULIN) COMPLEX (IMMUNORECEPTOR/IMMUNOGLOBULIN)
889	1tnm		147	232	1.7e-16	-0.49	0.29		MUSCLE PROTEIN TITIN MODULE M5 (CONNECTIN) 1TNM 3 (NMR, MINIMIZED AVERAGE	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									STRUCTURE) 1TNM 4 1TNM 58	
890	1f88	B	44	404	3.6e-65	0.01	-0.07		RHODOPSIN; CHAIN: A, B	SIGNALING PROTEIN PHOTORECEPTOR, G PROTEIN-COUPLED RECEPTOR, MEMBRANE PROTEIN, 2 RETINAL PROTEIN, VISUAL PIGMENT
892	1dva	L	215	313	5.1e-09	0.10	-0.06		DES-GLA FACTOR VIIA (HEAVY CHAIN); CHAIN: H, I; DES-GLA FACTOR VIIA (LIGHT CHAIN); CHAIN: L, M; (DPN)-PHE-ARG; CHAIN: C, D; PEPTIDE E-76; CHAIN: X, Y;	HYDROLASE/HYDROLASE INHIBITOR PROTEIN-PEPTIDE COMPLEX
892	1emn		48	123	6.8e-09	0.10	-0.20		FIBRILLIN; CHAIN: NULL;	MATRIX PROTEIN EXTRACELLULAR MATRIX, CALCIUM-BINDING, GLYCOPROTEIN, 2 REPEAT, SIGNAL, MULTIGENE FAMILY, DISEASE MUTATION, 3 EGF-LIKE DOMAIN, HUMAN FIBRILLIN-1 FRAGMENT, MATRIX PROTEIN
892	1f5y	A	29	105	1.7e-09	0.10	-0.07		LOW-DENSITY LIPOPROTEIN RECEPTOR; CHAIN: A;	LIPID BINDING PROTEIN LDL RECEPTOR; BETA HAIRPIN, 3-10 HELIX, CALCIUM BINDING
892	1fak	L	215	313	5.1e-09	-0.03	0.03		BLOOD COAGULATION FACTOR VIIA; CHAIN: L; BLOOD COAGULATION FACTOR VIIA; CHAIN: H; SOLUBLE TISSUE FACTOR; CHAIN: T; 5L15; CHAIN: I;	BLOOD CLOTTING COMPLEX(SERINE PROTEASE/COFACTOR/LIGAND), BLOOD COAGULATION, 2 SERINE PROTEASE, COMPLEX, CO-FACTOR, RECEPTOR ENZYME, 3 INHIBITOR, GLA, EGF, COMPLEX (SERINE 4 PROTEASE/COFACTOR/LIGAND), BLOOD CLOTTING
892	4mt2		179	236	3.4e-08	0.13	-0.19		METALLOTHIONEIN METALLOTHIONEIN ISOFORM II 4MT2 3	
892	9wga	A	140	292	6.8e-14	0.04	-0.19		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
892	9wga	A	172	304	1.4e-09	0.03	-0.19		LECTIN (AGGLUTININ) WHEAT	

SEQ ID NO:	PDB ID	Chain ID	Start AA	End AA	PSI-BLAST	Verify score	PMF score	SeqFold score	Compound	PDB annotation
									GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
892	9wga	A	3	185	3.4e-19	0.22	-0.19		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	
892	9wga	A	74	253	8.5e-18	0.14	-0.05		LECTIN (AGGLUTININ) WHEAT GERM AGGLUTININ (ISOLECTIN 2) 9WGA 3	

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE 6

SEQ ID NO:	Position of Signal Peptide	Maximum Score	Mean Score
447	1-18	0.984	0.928
448	1-30	0.937	0.671
450	1-26	0.976	0.902
452	1-21	0.973	0.927
453	1-16	0.881	0.748
459	1-47	0.981	0.720
461	1-40	0.957	0.708
464	1-26	0.908	0.748
465	1-15	0.986	0.828
467	1-18	0.986	0.971
468	1-19	0.916	0.649
469	1-27	0.954	0.804
470	1-37	0.992	0.827
471	1-17	0.949	0.860
472	1-35	0.978	0.702
473	1-35	0.990	0.881
474	1-47	0.990	0.833
477	1-19	0.966	0.845
479	1-20	0.944	0.721
504	1-30	0.937	0.671
523	1-26	0.976	0.902
527	1-23	0.978	0.911
536	1-26	0.982	0.944
564	1-21	0.973	0.927
565	1-16	0.881	0.748
600	1-21	0.985	0.885
645	1-47	0.981	0.720
647	1-23	0.975	0.886
698	1-26	0.908	0.748
702	1-25	0.972	0.930
703	1-35	0.974	0.788
706	1-37	0.969	0.747
715	1-15	0.986	0.828
731	1-18	0.986	0.971
732	1-20	0.978	0.824
742	1-19	0.916	0.649
743	1-13	0.956	0.798
748	1-27	0.954	0.804
766	1-17	0.949	0.860
786	1-35	0.978	0.702
797	1-17	0.989	0.926
805	1-32	0.980	0.785
815	1-47	0.990	0.833
836	1-48	0.969	0.712
840	1-22	0.997	0.951
845	1-19	0.953	0.798
856	1-43	0.973	0.682
858	1-23	0.974	0.873
867	1-25	0.988	0.888
889	1-16	0.964	0.890
891	1-19	0.966	0.845

TABLE 7

SEQ ID NO:	Chromosomal Location
1	2
2	22q12
3	12
4	12
5	13
6	15
7	15
8	11
9	1
10	22cen-q12.3
11	19
12	14
13	19
14	3
15	17q21.3-q22
16	22q13.2
17	22q13.2
18	16
19	11
20	14q32.33
21	14q32.33
22	22cen-q12.3
24	14q32.1
25	14
26	4
27	17
28	16
29	16
30	6p12
31	10
32	17
33	6
34	8
35	5q35.3
36	17p12-p11.2
37	12
38	18p11.23-p11.21
39	8
40	10
42	12p13
43	8p11
44	12
45	7p15-p14
46	11q24
47	2p24.1
49	10
50	10
51	10pter-q26.12
52	19
53	5
54	4
55	14
56	17
57	19q13.2

SEQ ID NO:	Chromosomal Location
58	20q13.12-13.13
59	2p23.3-q14.3
60	19
61	11
62	19q13.4
64	4q31.2-q31.3
66	12
67	13q34
68	12p11
69	1
70	17
71	5
72	14q11.2
73	19p13.1
74	14q
75	3
76	5
77	17q22-q24
78	2
80	1p36.3-p36.2
81	17
82	15
83	17
84	15
86	11p13
87	11p13
88	2p23.3-q34
89	6
90	Xq22
91	15q11.2
92	15q11.2
93	14
94	2p23.3-q31.1
95	6p21.2-21.3
96	4
97	5
98	16
99	9
100	1p32-p31
101	6
102	2p23.3-q14.3
103	6q14.3-q15
104	6q14.3-q15
105	19p12
106	16
107	1
108	2
109	18
110	3p21.1-9
111	17
112	20pter-p12.3
113	11q14
114	15
115	3
116	12q13
117	8pter-8p23.3

SEQ ID NO:	Chromosomal Location
118	4q34-q35
119	21q21
120	X
121	12q13
122	1p
123	16p13.1
124	17
125	10cen-q26.11
126	11
127	20p13-p12
128	2p11.2
129	4q32.1-q32.3
130	4
131	14
132	6q14.2-q16.1
133	3
134	8
135	19
136	1q25
138	17p13.1
139	12
140	15
141	9
142	4
143	12
144	20
145	21q22.11
146	11
147	7
149	7
150	X
151	15
152	19
153	2
154	7p21-p22
155	7q21.2-q31.1
157	1q21-q23
158	14
159	17q21
160	7q31-q32
161	12q22
162	14q21.1-q21.3
163	1
164	6
165	17
166	1
167	2
168	10q24.3
169	15
170	x
171	8
172	18
173	3
175	18
176	3
177	1

SEQ ID NO:	Chromosomal Location
178	2q12-q21
179	16
180	17
181	11q12-q13
182	16
183	17
184	15q11.2
185	12
186	9q32-q34.1
187	17q11-q21.3
188	1
189	1p34.1-p32
190	10
191	6p21
192	16
193	2q24.2
194	4
195	10cen-q26.11
196	4q31.2-q31.3
197	8q24.3
198	3q26.1-q26.2
199	16p13.3
201	6p11.2-p21.1
202	17q21
203	2
204	7
205	5
206	14q11-q12
207	22q13.1-q13.2
208	20q13.2-q13.3
209	17
210	1p36.11-36.23
211	5
212	3
214	17
216	11
217	1
218	1
219	Xp11.21-Xp11.23
220	12
221	2
222	11cen-11q12.3
223	6
224	6
225	15
226	22q13.1
227	22q13.1
228	22q12
229	14
230	1q25-q31
231	4q25-q27
232	14
233	3p25.3-3p24.1
234	3p25.3-3p23
235	15
237	17

SEQ ID NO:	Chromosomal Location
238	X
239	5q31.1
240	17
242	5q
243	8q22.2-q23
244	22
245	14
246	17
247	17
248	22q12.1
249	14
250	9
251	11q24-q25
252	17q12
253	17
255	2p24.3-p24.1
256	3p21.3
257	21q22.3
258	19p13-q13.4
259	11
260	Xq13.1
261	6
262	17
263	12q
264	4p16.1-p14
265	10p11.2
266	4
267	2p12
268	11cen-q12.1
269	3
270	11
271	8
272	11p15.3
273	11p15.3
274	11p15.3
275	2p23.3-q21.3
276	18
277	7
278	10q22.3-q23.2
279	10q22.3-q23.2
280	8p22
282	19
283	17
284	3
285	6p21.3
286	14q11.2
287	1
288	15
289	10cen-q26.11
290	22
291	4
292	1
293	1
294	4
296	1
297	1

SEQ ID NO:	Chromosomal Location
299	19
300	4q13-q21
302	22q11.2
303	9
304	3p13-q26.1
305	6q22.2-q22.33
306	17
307	17
308	19p13.1
309	17
310	12
311	17
313	3
314	15
315	15
316	14
317	10
318	2p24.3-p24.1
319	17
320	4
321	5q14
323	9
324	3p21
325	1
326	q13.1-13.2
327	17
329	2p14-p13
330	19pter-q12
331	20q11.1-q11.2
332	10
333	6q15-q16.1
335	11q11
336	22
337	7p13-p11.2
338	12q13
339	11p15.5-p15.4
340	4
341	11p15.5
342	7
343	22q12.1-q12.3
344	12q12-q13
345	18
346	16
347	20
349	12
350	4
351	6
352	19
353	17
354	15
355	3
356	14q24.3
357	19
358	11q13.5-q14.1
359	Xq25-26.1
360	19

SEQ ID NO:	Chromosomal Location
363	19
364	5
365	11p15
366	17
367	10
368	17
369	14q32
370	1
372	11
374	11q13
375	17
376	16
377	2
378	6
379	21p11
380	X
381	17
383	1q21
384	17
386	2
389	11p15.5
390	19
391	4
392	7p15.3-p21
394	20q13.3
395	4
397	14q11.2
398	4
399	14q11.2
400	22q13.1-q13.2
401	16
402	22q13.2-q13.3
406	22q11.23
407	15
408	3q27
409	22q12.2-13.1
411	16
413	2
415	12
417	7q21
418	11q23
420	12q12-q13
421	14
422	X
423	12
424	3q21
425	21q21.1-q21.2
429	15
430	9q34.3
431	x
432	2
433	18p11.23-p11.21
434	16
435	16
436	17
439	6

SEQ ID NO:	Chromosomal Location
440	5q14
441	2q33-q34
442	17
443	Xq22.2-q22.3
444	5p15.1-p14
446	9q34

TABLE 8

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Peptide Sequence	SEQ ID NO: in Priority Application USSN 09/687,527
1	447	6
2	448	7
3	449	8
4	450	9
5	451	12
6	452	13
7	453	14
8	454	15
9	455	16
10	456	18
11	457	20
12	458	21
13	459	22
14	460	23
15	461	24
16	462	26
17	463	27
18	464	28
19	465	29
20	466	30
21	467	31
22	468	32
23	469	33
24	470	34
25	471	35
26	472	37
27	473	38
28	474	40
29	475	41
30	476	46
31	477	48
32	478	49
33	479	50
34	480	51
35	481	52
36	482	53
37	483	54
38	484	55
39	485	56
40	486	57
41	487	58
42	488	59
43	489	60
44	490	61
45	491	62
46	492	63
47	493	64
48	494	65
49	495	66
50	496	67
51	497	68
52	498	69
53	499	70

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Peptide Sequence	SEQ ID NO: in Priority Application USSN 09/687,527
54	500	71
55	501	72
56	502	73
57	503	74
58	504	75
59	505	76
60	506	77
61	507	78
62	508	79
63	509	80
64	510	82
65	511	83
66	512	84
67	513	85
68	514	86
69	515	88
70	516	89
71	517	90
72	518	91
73	519	92
74	520	93
75	521	94
76	522	95
77	523	96
78	524	97
79	525	98
80	526	99
81	527	100
82	528	101
83	529	102
84	530	103
85	531	104
86	532	105
87	533	106
88	534	107
89	535	108
90	536	109
91	537	110
92	538	111
93	539	112
94	540	113
95	541	114
96	542	115
97	543	116
98	544	117
99	545	118
100	546	119
101	547	120
102	548	121
103	549	124
104	550	125
105	551	126
106	552	127
107	553	128
108	554	129

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Peptide Sequence	SEQ ID NO: in Priority Application USSN 09/687,527
109	555	130
110	556	131
111	557	132
112	558	133
113	559	134
114	560	135
115	561	136
116	562	137
117	563	138
118	564	139
119	565	140
120	566	141
121	567	142
122	568	143
123	569	144
124	570	146
125	571	147
126	572	148
127	573	149
128	574	150
129	575	151
130	576	152
131	577	153
132	578	154
133	579	155
134	580	156
135	581	157
136	582	158
137	583	160
138	584	161
139	585	162
140	586	163
141	587	164
142	588	165
143	589	166
144	590	167
145	591	168
146	592	169
147	593	170
148	594	171
149	595	172
150	596	173
151	597	174
152	598	175
153	599	176
154	600	177
155	601	178
156	602	179
157	603	180
158	604	181
159	605	182
160	606	183
161	607	184
162	608	185
163	609	186

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Peptide Sequence	SEQ ID NO: in Priority Application USSN 09/687,527
164	610	187
165	611	188
166	612	189
167	613	190
168	614	191
169	615	192
170	616	193
171	617	194
172	618	196
173	619	197
174	620	198
175	621	199
176	622	200
177	623	201
178	624	202
179	625	203
180	626	205
181	627	206
182	628	207
183	629	208
184	630	209
185	631	210
186	632	211
187	633	212
188	634	213
189	635	214
190	636	215
191	637	216
192	638	217
193	639	218
194	640	219
195	641	220
196	642	221
197	643	223
198	644	224
199	645	225
200	646	226
201	647	227
202	648	228
203	649	229
204	650	230
205	651	231
206	652	232
207	653	233
208	654	234
209	655	235
210	656	236
211	657	237
212	658	238
213	659	240
214	660	241
215	661	243
216	662	244
217	663	245
218	664	246

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Peptide Sequence	SEQ ID NO: in Priority Application USSN 09/687,527
219	665	247
220	666	248
221	667	249
222	668	250
223	669	251
224	670	252
225	671	253
226	672	254
227	673	255
228	674	256
229	675	257
230	676	258
231	677	259
232	678	260
233	679	261
234	680	262
235	681	263
236	682	264
237	683	265
238	684	266
239	685	267
240	686	268
241	687	269
242	688	270
243	689	271
244	690	272
245	691	273
246	692	274
247	693	275
248	694	276
249	695	277
250	696	278
251	697	279
252	698	280
253	699	281
254	700	282
255	701	283
256	702	284
257	703	285
258	704	286
259	705	287
260	706	288
261	707	289
262	708	290
263	709	291
264	710	292
265	711	293
266	712	294
267	713	296
268	714	297
269	715	298
270	716	299
271	717	300
272	718	301
273	719	302

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Peptide Sequence	SEQ ID NO: in Priority Application USSN 09/687,527
274	720	303
275	721	304
276	722	305
277	723	306
278	724	307
279	725	308
280	726	309
281	727	310
282	728	312
283	729	314
284	730	315
285	731	316
286	732	317
287	733	318
288	734	319
289	735	320
290	736	321
291	737	322
292	738	323
293	739	324
294	740	325
295	741	326
296	742	327
297	743	328
298	744	329
299	745	330
300	746	331
301	747	332
302	748	333
303	749	334
304	750	335
305	751	337
306	752	338
307	753	339
308	754	340
309	755	341
310	756	342
311	757	343
312	758	344
313	759	345
314	760	346
315	761	347
316	762	348
317	763	350
318	764	351
319	765	352
320	766	353
321	767	354
322	768	355
323	769	356
324	770	357
325	771	358
326	772	359
327	773	360
328	774	361

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Peptide Sequence	SEQ ID NO: in Priority Application USSN 09/687,527
329	775	362
330	776	363
331	777	364
332	778	365
333	779	366
334	780	367
335	781	368
336	782	369
337	783	370
338	784	371
339	785	372
340	786	373
341	787	375
342	788	376
343	789	377
344	790	378
345	791	379
346	792	380
347	793	381
348	794	382
349	795	383
350	796	384
351	797	385
352	798	386
353	799	387
354	800	388
355	801	389
356	802	390
357	803	391
358	804	392
359	805	393
360	806	394
361	807	395
362	808	396
363	809	397
364	810	398
365	811	399
366	812	400
367	813	401
368	814	402
369	815	403
370	816	404
371	817	405
372	818	406
373	819	407
374	820	408
375	821	409
376	822	410
377	823	411
378	824	412
379	825	413
380	826	414
381	827	415
382	828	416
383	829	417

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Peptide Sequence	SEQ ID NO: in Priority Application USSN 09/687,527
384	830	418
385	831	419
386	832	420
387	833	421
388	834	422
389	835	423
390	836	424
391	837	425
392	838	426
393	839	427
394	840	428
395	841	429
396	842	430
397	843	431
398	844	432
399	845	433
400	846	434
401	847	435
402	848	436
403	849	438
404	850	439
405	851	440
406	852	441
407	853	442
408	854	443
409	855	444
410	856	445
411	857	446
412	858	447
413	859	448
414	860	449
415	861	450
416	862	451
417	863	452
418	864	453
419	865	454
420	866	455
421	867	456
422	868	457
423	869	458
424	870	459
425	871	460
426	872	461
427	873	462
428	874	463
429	875	464
430	876	465
431	877	467
432	878	468
433	879	469
434	880	470
435	881	471
436	882	472
437	883	473
438	884	474

SEQ ID NO: of Full-length Nucleotide Sequence	SEQ ID NO: of Full-length Peptide Sequence	SEQ ID NO: in Priority Application USSN 09/687,527
439	885	475
440	886	476
441	887	477
442	888	478
443	889	479
444	890	480
445	891	481
446	892	482

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-446, a mature protein coding portion of SEQ ID NO: 1-446, an active domain coding portion of SEQ ID NO: 1-446, and complementary sequences thereof.

2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.

3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.

4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.

5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.

6. A vector comprising the polynucleotide of claim 1.

7. An expression vector comprising the polynucleotide of claim 1.

8. A host cell genetically engineered to comprise the polynucleotide of claim 1.

9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.

10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:

(a) a polypeptide encoded by any one of the polynucleotides of claim 1; and

(b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO: 1-446.

11. A composition comprising the polypeptide of claim 10 and a carrier.

12. An antibody directed against the polypeptide of claim 10.

13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and

b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.

14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;

b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and

c) detecting said product and thereby the polynucleotide of claim 1 in the sample.

15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.

16. A method for detecting the polypeptide of claim 10 in a sample, comprising:

a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and

b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.

17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and

b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

19. A method of producing the polypeptide of claim 10, comprising,

a) culturing a host cell comprising a polynucleotide sequence selected from SEQ ID NO: 1-446, a mature protein coding portion of SEQ ID NO: 1-446, an active domain coding portion of SEQ ID NO: 1-446, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-446, under conditions sufficient to express the polypeptide in said cell; and

b) isolating the polypeptide from the cell culture or cells of step (a).

20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of any one of the polypeptides SEQ ID NO: 447-892, the mature protein portion thereof, or the active domain thereof.

21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.

22. A collection of polynucleotides, wherein the collection comprising the sequence information of at least one of SEQ ID NO: 1-446.

23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.

24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.

25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.

26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

Cell, Vol. 104, 531–543, February 23, 2001, Copyright ©2001 by Cell Press

Obesity and the Regulation of Energy Balance

Review

Bruce M. Spiegelman*‡ and Jeffrey S. Flier†‡
*Dana-Farber Cancer Institute and the Department of Cell Biology
Harvard Medical School
Boston, Massachusetts 02115
†Beth Israel Deaconess Medical Center
Division of Endocrinology and the Department of Medicine
Harvard Medical School
Boston, Massachusetts 02115

Obesity is defined medically as a state of increased body weight, more specifically adipose tissue, of sufficient magnitude to produce adverse health consequences. There has been an alarming increase recently in the prevalence of this heterogeneous group of disorders in the Western world (Kuczmarski et al., 1994). Fully one-third of the American population is now considered obese, and the prevalence of obesity in children is escalating dramatically, presaging even greater medical harm in the decades to come (Troiano and Flegal, 1999). What accounts for this epidemic of energy storage? Body weight and composition, and the storage of energy as triglyceride in adipose tissue, are determined by the interaction between genetic, environmental, and psychosocial factors. These influences ultimately act by changing the energy balance equation, that is, the long-term balance between energy intake and expenditure. Physiologic studies had previously suggested that body weight and energy stores are homeostatically regulated, with either weight loss or gain producing concerted changes in energy intake and expenditure that resist the initial perturbation. Recent cloning of several obesity genes has revealed the initial molecular components of a coherent physiologic system for energy homeostasis (Barsh et al., 2000). Studies of obesity pathogenesis must now attempt to explain the disorder in the context of this physiologic system.

Although the role of genes in body fat regulation is now established, it is safe to assume that the rising prevalence of obesity has not been due to a recent change in the genetics of the Western world. The propensity for obesity must have been in our midst for a long time, only to emerge recently on a large scale as a result of changes in the environment, in particular the availability and composition of food and reduced requirement for physical exertion. It is very likely that the ability to store fat in times of nutritional abundance was a positive trait selected over many thousands of years of human evolution. The idea that humans evolved to efficiently store excess energy as fat to deal with periodic famine has been given a name — the "thrifty gene" hypothesis (Neel, 1999). An obese human of approximately 250 lbs. has the energy stores to survive a total fast of approximately 150 days! This impressive energy reserve is due both to the high energy content of triglycerides versus polysaccharides, and the fact that triglycerides are stored in essentially anhydrous form; polysaccharides such as glycogen are hydrated in storage form, decreasing their efficiency as fuel.

To understand obesity, one must understand the concept of energy balance (Figure 1). Assuming that an individual has no problem with the absorption of nutrients, stored energy will increase *only* if energy intake exceeds total body energy expenditure. Energy expenditure takes the form of physical activity, basal metabolism, and adaptive thermogenesis. Physical activity refers to all voluntary movement, while basal metabolism refers to the myriad biochemical processes necessary to sustain life. Adaptive thermogenesis refers to energy dissipated in the form of heat in response to environmental changes, such as exposure to cold and alterations in diet. It should be pointed out that the boundary between what is considered basal metabolism versus adaptive thermogenesis is not always clear-cut. Mammals often live in climates with temperatures below body temperature, sometimes far below body temperature. Thus, the determination of energy expenditure in response to cold versus that which is considered part of basal metabolism can be quite arbitrary. Traditionally, the basal metabolism rate is defined as the energy expenditure of a subject relaxed and at rest, at thermoneutrality, 8–12 hr after the last food ingestion. Metabolic rates of mammals will obviously vary as a continuum depending upon precise environmental conditions to which the organism is exposed.

Various cellular events can generally cause or prevent obesity *only* if they affect the overall energy equation of the individual. For example, absent the concept of energy balance, it might be supposed that obesity and its complications could be ameliorated by directly limiting adipose tissue development. Experimental models in which fat cell differentiation, development, and survival have been directly reduced have shown that this is not the case. Far from creating healthy, lean mice, the resulting animals appear to suffer from lack of an appropriate depot in which to deposit excess energy, and loss of key adipocyte-derived hormones (Ross et al., 1993; Moitra et al., 1998; Shimomura et al., 1998). This results in increased levels of blood lipids, accumulation of fat in the liver, diabetes, and death. Conversely, when a genetically altered animal is found to be leaner, this *does not* necessarily imply that the targeted gene has a role in fat development, per se. Rather, a defect is created in some component in the total energy balance scheme—usually food intake or energy expenditure.

A precise understanding of the contribution of diet to obesity has been confounded by the difficulty of obtaining accurate measurements of food intake in free-living individuals. Obese individuals tend to underreport food intake by as much as 30% (Lichtman et al., 1992), and most obese individuals are believed to ingest more calories than lean individuals (matched for exercise and other features), to maintain their increased weights.

‡E-mail: bruce_spiegelman@dfci.harvard.edu (B. M. S.); jflier@caregroup.harvard.edu (J. S. F.).

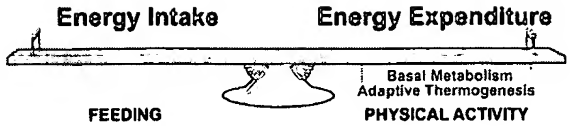

Figure 1. Key Component of the Energy Balance System

Obesity develops only if energy intake, in the form of feeding, chronically exceeds total body expenditure. Energy expenditure includes physical activity, basal metabolism, and adaptive thermogenesis.

Apart from the obvious effect of availability of palatable foods, the study of the influence of macronutrient composition, such as the balance of protein, carbohydrate, and fat on long-term body weight is in its infancy, and gene–diet interactions in the feeding response to diets of varying composition are just beginning to be studied. Because of the central notion of energy balance, food intake can be thought of as inadequate or excessive only in the context of that individual's energy expenditure.

It is clear that different individuals have a certain genetic propensity to store excessive caloric intake as fat. In a classic study, Bouchard and collaborators overfed pairs of monozygotic twins by precisely calibrated amounts (Bouchard et al., 1990). Different sets of twins showed remarkable differences in the degree to which these calories were stored as fat, but the tendency toward increased adiposity within each set of twins was remarkably similar. Since overfeeding above basal needs was controlled, the likely difference in fat accretion *between* sets of twins was likely due to differences in some component of energy expenditure, perhaps basal metabolic rates or adaptive thermogenesis. Further evidence that differences in metabolic rate are important variables in human obesity is provided by prospective studies done in Pima Indians (Ravussin, 1995), a group of Native Americans with a high predilection for obesity. When individuals were typed as to metabolic rates and then followed, those with lower metabolic rates had a greater incidence and magnitude of obesity.

CNS Control of Energy Intake and Body Weight

The central nervous system (CNS) influences energy balance and body weight through three mechanisms: (1) effects on behavior, including feeding and physical activity; (2) effects on autonomic nervous system activity, which regulates energy expenditure and other aspects of metabolism; and (3) effects on the neuroendocrine system, including secretion of hormones such as growth hormone, thyroid, cortisol, insulin, and sex steroids. The identity and coordination of these complex systems has been the subject of intense study, and much recent progress.

Regulation of Energy Intake: Short- and Long-Term Control

Feeding behavior lies at the interface between free will and physiology, and is influenced by many factors. In addition to food availability, feeding is affected by metabolic, neural, and endocrine factors, and is modified by powerful visual, olfactory, emotional, and cognitive inputs. Ultimately, all of these factors must be integrated, so that decisions to begin and end periods of feeding will result (Schwartz et al., 2000). The regulation of feeding behavior may be divided into short- and long-term control systems. Short-term control involves the initiation and termination of meals. The major determinant of meal size is the onset of satiety, a response to neural and endocrine factors, such as gut distension and release of the gut peptide cholecystokinin (CCK) that are generated during the course of meal ingestion (Moran, 2000). These signals are transmitted to the caudal brainstem via the vagus nerve, where integration with other inputs occurs, leading to meal termination. Regulation of individual meal size by factors induced by meal ingestion is insufficient to account for energy balance over long periods of time since mice repeatedly injected with CCK maintain weight by ingesting more meals of smaller size. Long-term signals that reflect the status of energy stores, such as the fat-derived hormone leptin, provide information to the CNS that further regulates feeding behavior to promote energy homeostasis. Not surprisingly, these short- and long-term systems are interrelated (Emond et al., 1999), such that the feeding response to energy deficit is accomplished predominantly through increased meal size.

Role of the Hypothalamus

The hypothalamus is a region of the brain critical for regulation of homeostatic processes such as feeding, thermoregulation, and reproduction (Elmquist et al., 1999). To accomplish these ends, the hypothalamus senses neural, endocrine, and metabolic signals, integrates these inputs, and engages distinct effector pathways, resulting in behavioral, autonomic, and endocrine responses. In addition to the hypothalamus, central control of appetite and energy balance clearly involves widely distributed neural systems in the brainstem, cerebral cortex, olfactory areas, and elsewhere. The central role of the hypothalamus in appetite and satiety was determined early on by lesion studies. Lesions in the ventromedial hypothalamus cause obesity, while lesions in the lateral hypothalamus cause leanness (Elmquist et al., 1999). The inherent crudeness of physical lesions and the absence of defined molecular and anatomic pathways underlying them limited the interpretation of this paradigm, however, and other approaches were taken. For example, the effects of centrally administered neurotransmitters such as norepinephrine, dopamine, and serotonin, and hypothalamic neuropeptides such as NPY and CRH on food intake, autonomic output, metabolism, and energy balance were assessed. These studies established the existence of several CNS ligand-receptor pathways capable of modifying energy intake, energy balance, and metabolic status.

Early Studies of NPY

Examination of the neuropeptide Y (NPY) pathway best exemplifies such studies. NPY is widely and abundantly expressed within the nervous system. The arcuate nucleus of the hypothalamus is one site of particularly dense expression, and nutritional regulation (increased with starvation) is uniquely observed at that site. When administered into cerebral ventricles or specific hypothalamic nuclei, NPY robustly and rapidly increases feeding and suppresses energy expenditure, and thereby promotes obesity (Stanley et al., 1986; Billington et al., 1994). NPY was therefore viewed as an excellent candidate for an endogenous regulator of energy balance, promoting anabolism in response to energy deficits. The

biggest gap in knowledge regarding NPY, and the many other neurochemicals similarly shown to stimulate or suppress feeding, was our limited understanding of the physiologic system in which these powerful neural circuits participated. Specifically, how did nutritional status communicate information to these central pathways for integration and control? One obvious candidate signal was the hormone insulin (Schwartz et al., 2000). Insulin levels do reflect energy balance and stores, as they fall with starvation and rise with obesity. Insulin is transported into the brain through a saturable process, and when centrally administered, suppresses both food intake and arcuate NPY expression. However, the absence of insulin, as in type 1 diabetes, is associated with weight loss rather than gain. Furthermore, evidence that physiologic changes in peripheral insulin levels affected energy balance through the brain was limited, so most investigators believed that insulin was not the dominant peripheral signal to the CNS for regulation of energy balance. That signal was yet to be discovered.

Monogenic Obesities and the Discovery of Leptin

The field of energy balance advanced rapidly with the cloning of genetic loci responsible for several of the previously identified monogenic obesity syndromes in mice. The most dramatic and consequential advance resulted from the identification of the basis for syndromes of obesity in ob/ob and db/db mice (Zhang et al., 1994; Chen et al., 1996). Prior studies used the technique of parabiosis, in which mice are surgically joined to permit passage of molecules from one to the other. These led to the suggestion 30 years ago that ob/ob mice might be deficient in a circulating signal of satiety, while db/db mice might be deficient in its cognate receptor. This prescient prediction was fully born out with cloning of these genes in 1994 and 1995. The ob gene encodes a unique member of the cytokine family now named leptin, from the Greek root leptos, for thin. The dominant site from which leptin is secreted is the adipocyte, and the protein is truncated and biologically inactive in mutant mice (Zhang et al., 1994). The obesity syndrome in ob/ob mice is corrected by administration of the missing hormone (Halaas et al., 1995; Friedman and Halaas, 1998). Regulated expression of leptin in other sites, such as skeletal muscle (Wang et al., 1998), placenta, and stomach has been reported, and may ultimately be proven to be physiologically important, although this is not yet established. The signaling form of the leptin receptor, ObRb, is deleted in db/db mice (Chen et al., 1996; Friedman and Halaas, 1998), which are consequently unresponsive to endogenous or exogenous leptin. The identification of these two proteins establishes the first well-documented components of a powerful nutritional feedback loop from adipose tissue to the brain.

Leptin Physiology

Despite the remarkable ability of leptin to reverse obesity in leptin deficient ob/ob mice, and to cause leanness in wild-type mice, the function of the leptin pathway may not be simply understood as an antiobesity axis. Indeed, substantial data suggest that basal levels of leptin in the fed state serve as a signal of energy sufficiency (Reviewed in Ahima and Flier, 2000). Withdrawal of the leptin signal occurs quite rapidly with food restriction, exceeding the rate at which fat stores are reduced (Ahima et al., 1996). Reduced leptin entrains a complex neural response characteristic of starvation that includes hunger/food seeking behavior, efficient metabolism (clearly demonstrated in rodents), and an array of neuroendocrine responses that favor survival during periods of limited energy, such as suppression of reproduction, linear growth, and thyroid hormone levels (Ahima et al., 1996). Reintroduction of energy supplies rapidly raises leptin levels and suppresses this starvation program. Absence of the leptin signal in the presence of sufficient energy promotes obesity in both rodents and humans by producing an internal perception of starvation in the midst of plenty. In addition to its actions through CNS circuits, leptin appears to exert several effects directly on peripheral tissues through leptin receptor signaling. One important effect may involve suppression of triglyceride accumulation in nonadipose tissue, such as muscle and liver, which contributes to insulin resistance (Lee et al., 2000). Potent effects of leptin have also been seen on the immune system, the vascular system, and even on bone turnover (Ducy et al., 2000). Although it clearly serves as the switch from the starved to the fed states, leptin has limits in controlling obesity. As fat mass increases, further rises in leptin have a limited ability to suppress food intake and prevent obesity, as seen by the prevalence of obesity despite high levels of circulating leptin (Considine et al., 1996). Thus, the antiobesity role of leptin might have been limited through evolutionary pressure to promote fat storage in times of plenty.

Central Neural Circuits Regulating Energy Balance

Knowledge of the neural circuits that coordinate the response to leptin and other inputs has increased at a rapid rate. Through positional cloning of rodent genes, targeted gene deletion, identification of mutant genes in human obesity, and follow-up studies using the techniques of functional neuroanatomy, a complex and extensive central circuit for regulation of energy balance has been defined. Although several sites in the brain express leptin receptors and respond to this hormone and various neuropeptides with changes in energy intake and expenditure, the best characterized and most clinically relevant circuit is simply described as a leptin-regulated central melanocortin circuit. The major components of this circuit are presented in Figure 2, where 7 unique proteins in a single pathway have been shown to contribute to weight regulation and obesity.

The Central Melanocortin Pathway

Leptin acts through ObRb receptors on two distinct populations of neurons in the arcuate nucleus. One population coexpresses the orexigenic (feeding-inducing) neuropeptides NPY and AgRP, and leptin action reduces their expression (Elias et al., 1999; Elmquist et al., 1999; Schwartz et al., 2000). The other population coexpresses mRNAs encoding anorexigenic peptides, cocaine and amphetamine related transcript (CART) and α-MSH (derived from proopiomelanocortin [POMC]), and leptin induces their expression (Elias et al., 1999; Elmquist et al., 1999; Schwartz et al., 2000). Thus, leptin suppresses two orexigenic peptides and induces two anorexigenic peptides through direct action on arcuate neurons. The pathway just described is incestuous, as AgRP and α-MSH are antagonistic ligands for a common receptor,

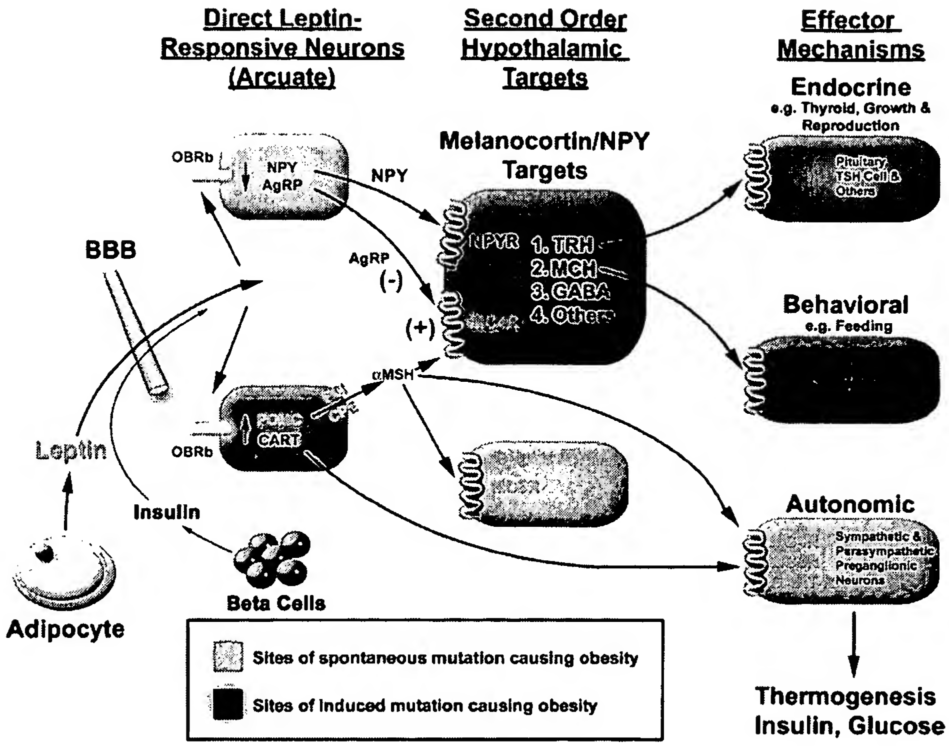

Figure 2. A Leptin-Regulated Melanocortin Circuit Influences Energy Homeostasis and Body Weight

The adipocyte hormone leptin crosses the blood brain barrier (BBB) and acts directly on two populations of neurons within the arcuate nucleus that express NPY and AgRP or POMC and CART. Leptin stimulates production of α-MSH, an agonist for the MC4 receptor (as well as CART), and inhibits production of AgRP, an antagonist for this receptor (as well as NPY). MC4 receptor–expressing neurons receive these leptin-regulated signals, as well as others, such as NPY. Such MC4R neurons are just now being chemically and functionally identified, and include TRH neurons in the paraventricular nucleus (PVH) that regulate the thyroid, MCH neurons in the lateral hypothalamus that regulate feeding, GABAergic neurons in the PVH that modify other as yet unidentified neurons tied into energy balance, and others. Several outputs of the MC4R-expressing neurons include: endocrine outputs such as thyroid, growth and reproduction, through control of pituitary function; behavioral outputs, including feeding; autonomic output, regulating energy expenditure; insulin secretion; and glucose homeostasis. Sites in the pathway at which spontaneous loss of function mutations have caused obesity in rodents and humans are indicated in yellow, as are sites at which induced mutations have caused obesity in rodents (in blue). Not shown here are potential direct actions of leptin on peripheral tissues.

the melanocortin 4 receptor (MC4R), which is expressed primarily in the brain (Cone, 1999). Activation of MC4R by MSH reduces food intake, while suppression of MC4R signaling through this receptor by the endogenous antagonist AgRP or pharmacologic antagonists increases feeding and diminishes the hypophagic response to leptin (Fan et al., 1997). This pathway was discovered through the convergence of several prior lines of investigation. The dominant obesity syndrome of the A^y mouse was shown to be due to a gene rearrangement causing ectopic expression of the coat color regulating protein agouti (Bultman et al., 1992). Through transgenic and pharmacologic experiments, it became apparent that agouti produced obesity by antagonizing the action of α-MSH on MCR4s within the brain (Lu et al., 1994; Fan et al., 1997; Ollmann et al., 1997; Graham et al., 1997), thereby mimicking the hypothalmic agouti homolog AgRP.

Gene deletion of the MC4R causes obesity in mice, and mice heterozygous for the knockout allele have moderate obesity as well (Huszar et al., 1997). Remarkably, 4%–5% of severe human obesity appears to be due to mutation at this locus, and most affected humans have a single mutant allele, which causes obesity through haploinsufficiency, rather than a dominant-negative mechanism (Farooqi et al., 2000). This suggests that this pathway is required for normal energy homeostasis and is extremely tightly regulated. The obesity in several other rare human and murine syndromes also converges on this pathway. For example, mutation in the POMC gene, which prevents production of POMC products including α-MSH, produces obesity in mice and humans (Krude et al., 1998). Likewise, mutation in neuropeptide processing enzymes PC-1 and carboxypeptidase E cause complex obesity syndromes in humans and mice respectively, very likely at least in part through effects on POMC processing (Reviewed in Barsh et al., 2000). Very recently, targeted deletion of the MC3R, a closely related receptor also restricted largely to the brain, was shown to produce obesity in mice (Butler et al., 2000). Interestingly, the obesity from this lesion occurs without the hyperphagia seen in MC4R mutants, and may be associated with a loss of lean body mass as well as increased adiposity. Thus, these two melanocortin receptors can cause obesity through distinct physiologic mechanisms.

Pathways Downstream of Melanocortin Receptors

The identification of a leptin-regulated melanocortin pathway provides a molecular and neuroanatomic link

between peripheral signals and CNS circuits, but leaves open the question of how these melanocortin signals produce downstream effects on appetite, energy expenditure, and neuroendocrine function. Several possible mechanisms are emerging. In one model, leptin-regulated arcuate melanocortin nerve terminals project onto neurons within the paraventricular hypothalamic nucleus (PVN) that have previously been described to respond in vivo to changes in nutritional status and leptin levels. The PVN can be viewed as a motor arm of the hypothalamus, as it regulates pituitary hormone secretion via specific neuropeptides released by projections to the median eminence, and regulates autonomic activity via projections to autonomic preganglionic neurons. One example is the TRH neuron in the PVN that regulates the pituitary–thyroid axis. TRH expression within the PVN is regulated through both melanocortin inputs from the arcuate nucleus acting through MC4Rs on TRH neurons, and by direct action via leptin receptors on these cells (reviewed in Flier et al., 2000). Similar mechanisms may account for leptin effects on other PVN neurons that influence endocrine status, autonomic function, or appetite. A second and parallel model involves direct projection of these arcuate melanocortinergic neurons (AgRP and α-MSH) onto neurons within the lateral hypothalamus that express the orexigenic neuropeptides MCH and orexin/hypocretin (Elias et al., 1999). MCH, the expression of which was discovered to be upregulated in ob/ob hypothalami, stimulates food intake (Qu et al., 1996). Deletion of the MCH gene causes a lean phenotype (Shimada et al., 1998), and transgenic overexpression promotes obesity. The first two models propose direct actions of melanocortins on TRH or MCH neurons. Additional mechanisms are likely to exist, since electrophysiologic evidence supports melanocortinergic neurons projecting to GABAergic interneurons in the PVN that are proposed to serve as integrators of numerous inputs (Cowley et al., 1999). It is likely that all of these mechanisms exist in concert.

Other Neuropeptides and Neurotransmitters

The melanocortin system, although extremely important, is not the only neuropeptide system involved in weight regulation. NPY acts through several species of GPCRs to regulate energy balance, and PVN neurons (either interneurons or neurons such as TRH neurons) may be sites where melanocortin and NPY signals are integrated. Although $NPY^{-/-}$ mice feed normally and have normal body weight, NPY deficiency ameliorates obesity and other features of ob/ob mice, indicating that NPY is necessary for the full response to leptin deficiency (Erickson et al., 1996a, 1996b). CART is expressed widely, and is coexpressed in many leptin-regulated arcuate POMC neurons. Among several projections, leptin regulated CART neurons in the arcuate nucleus project to autonomic sites in the spinal cord, providing a possible link to autonomic pathways (Elias et al., 1998). Many other neuropeptides, including CRH, GHRH, and galanin have been described to participate in these regulatory pathways, and most are beyond the scope of this review. Ghrelin is a peptide expressed in stomach and brain that was originally identified through its actions on the growth hormone axis. It is now clear that ghrelin promotes hyperphagia and obesity through actions in the brain, possibly on NPY neurons (Tschop et al., 2000). It is likely that additional peptides will be identified. In addition to neuropeptides and transmitters, the function of these circuits is also influenced by metabolic fuels. Neurons in the hypothalamus that respond to changes in glucose levels (low brain glucose promotes feeding) may be the same as, or functionally linked to those neurons that respond to leptin and express the peptides discussed above. A role for lipid mediators in metabolic sensing may be suggested by the recent observation that inhibitors of fatty acid synthase potently inhibit food intake through actions in the brain (Loftus et al., 2000).

The neurotransmitters norepinephrine, dopamine and serotonin are well known to be involved in central energy balance circuits. Serotonergic neurons within the caudal brainstem project widely within the brain, and drugs that increase serotonergic signaling suppress food intake and have been used to treat obesity. Mice with deletion of the $5HT_{2c}$ serotonin receptor subtype have modest obesity (Nonogaki et al., 1998). Leptin increases serotonin turnover, suggesting that these pathways can converge, but $5HT_{2C}$-deficient mice retain an anorectic response to leptin.

The Return of Insulin

Although the discovery of leptin overshadowed earlier interest in the role of insulin as a central regulator of energy balance, the venerable hormone has made a comeback. Not only is insulin a prominent positive regulator of leptin expression in the fat cell, but the mild obesity in mice with neuron-specific deletion of the insulin receptor (Bruning et al., 2000) or the insulin receptor substrate IRS-2 (Burks et al., 2000), support the idea that insulin and leptin may cooperate in central pathway regulation. It will be important to determine whether leptin and insulin signaling pathways converge on some of the same target cells, and by what mechanisms these signaling pathways might interact.

Leptin Resistance: Hard Wired or Acquired?

As discussed above, most obese humans and rodents develop obesity despite high leptin levels, and administration of additional leptin fails to reverse the obese state. It is possible that "leptin resistance" arose through evolution to permit energy storage in times of plenty. This still leaves open the question of mechanism. One potential mechanism involves a limitation of leptin transport across the blood brain barrier, which may operate through alternative splice variants of the leptin receptor highly expressed in brain microvessels. This mechanism is supported by the ability of leptin injected directly into the brain to both suppress food intake and induce hypothalamic signaling more effectively than leptin injected by the peripheral route in mice with diet-induced obesity (Van Heek et al., 1997; El-Haschimi et al., 2000). The molecular basis for this limitation is not yet defined. Leptin signaling within the hypothalamus may also be impaired in the obese state. In susceptible C57Bl mice with obesity induced by high fat diet, this signaling defect is acquired as obesity develops (El-Haschimi et al., 2000). Leptin resistance in these common states is not complete, however, since obesity is much less severe than seen in states of absent ligand or receptor, and these mice lack several features, such as neuroendocrine defects, that characterize mice completely lacking leptin or leptin receptors. This would make good evolutionary sense, since modifications of leptin action designed to permit energy storage would be best designed

to avoid the drastic consequences of starvation that result from total leptin lack. Currently, we do not understand the details of leptin signaling and targets that result in such discordant effects.

Leptin Signaling and Resistance

Leptin receptors are members of the class I cytokine receptor family that utilize associated Jak kinases for signal transduction (Tartaglia, 1997). The best studied aspect of leptin signaling is the Jak-dependent activation of the STAT 3 pathway, and subsequent regulation of target gene expression. It is also clear that leptin also induces, via Jak, activation of the MAPkinase pathway, one mechanism for which involves the participation of the phosphatase SHP-2. It has not yet been determined whether distinct effects of leptin, including rapid effects on ion channel activity, require specific downstream effector mechanisms. The control of cytokine signaling involves negative feedback signals. SOCS-3 is a member of one such family of negative regulatory proteins that is induced in leptin responsive cells by leptin signaling pathways, and can serve as a marker of cells responding directly to leptin (Bjorbaek et al., 1998, 1999). SOCS-3 inhibits leptin signaling by actions at the level of Jak as well as through binding to the receptor itself (Bjorbaek et al., 2000). It is not yet clear whether endogenous SOCS-3, or other regulators such as the STAT inhibitors PIAS-1/3 are responsible for leptin resistance in obesity. Since leptin and insulin may have some common sites of action, it is interesting to note that insulin signaling may also be antagonized by SOCS family members (Emanuelli et al., 2000).

Genes and the Environment

The cloning of genes responsible for the previously known rodent monogenic obesities (e.g., ob, db, and A^y) has led to the identification of regulatory pathways, and has indicated that these pathways are preserved in humans. The realization that severe obesity in humans can result from mutations in the ob, db, and MC4R loci, with the latter accounting for 4%–5% of severe cases, indicates the importance of these systems. However, the rarity of these mutations highlights the fact that most human obesity is polygenic rather than Mendelian, controlled by many genetic loci. Since the prevalence of obesity is increasing in industrialized societies, it is apparent that many of these genes must confer susceptibility to environmental factors, such as availability of food and composition of diets, and response to exercise, or lack of it. Efforts are underway to map genes that confer susceptibility to diet induced obesity in inbred strains of mice. Responsible genes may be part of already identified pathways, or may be previously unknown components of known pathways. In human populations, genome scans for linkage with obesity-related phenotypes are ongoing in many populations. One locus showing linkage to leptin levels in two populations is on the short arm of chromosome 2 at band 21 (2p21), a region that includes the POMC gene (Comuzzie et al., 1997). In addition to studies based on linkage, or detection of mutations in rare Mendelian disorders, many studies have sought evidence through a candidate gene approach. Although this approach can be successful when the candidate has a fundamental role (e.g., the MC4R locus), the list of plausible candidates is very long, and factors such as heterogeneity of populations and small biological impact of a given variation have led to inconsistent findings (e.g., the W64R variation in the β3 adrenergic receptor) (Reviewed in Barsh et al., 2000).

Control of Energy Expenditure—Adaptive Thermogenesis

The components of energy expenditure that can be readily altered, i.e., physical activity and adaptive thermogenesis, are of particular interest in the control of obesity. Since physical activity is more properly the realm of fitness gurus and psychologists, we will concentrate here on describing the current state of scientific thinking about adaptive thermogenesis. Work in this area has centered on two main aspects: the neural circuitry that activates thermogenesis and the peripheral tissues that actually oxidize fuels.

One might have assumed that neural pathways that control food intake and energy expenditure are completely distinct. However, both naturally occurring and targeted genetic lesions in mice have indicated that these pathways are tightly interrelated. Mutations in leptin and the leptin receptor, MCR4 and MCH all strongly influence *both* food intake and energy expenditure in a coherent way, so that fat storage is increased (or decreased) via both of these major components of the energy balance equation. Conversely, exogenous administration of leptin to leptin-deficient mice decreases food intake, but also keeps energy expenditure at a higher level than would be expected for that degree of food intake (Friedman and Halaas, 1998). Clearly, a major component of this brain-driven thermogenesis is dependent on output from the sympathetic nervous system to the peripheral tissues, especially brown fat and skeletal muscle. While the role of the sympathetic nervous system in activating brown adipose tissue (BAT)-mediated thermogenesis via β-adrenergic receptors and cyclic AMP (cAMP) is well established, the role in skeletal muscle is still a matter of some conjecture. That catecholamines infused into skeletal muscle increase energy expenditure without the performance of work is quite clear (Simonsen et al., 1992); whether the key step(s) involves control of vascular tone, or regulation of specific components involved in fuel oxidation is not established. However, the fact that mice genetically ablated for the production of catecholamines show greater sensitivity to cold than do mice mutated in UCP-1 (Enerback et al., 1997; Thomas and Palmiter, 1997) (see below), suggests strongly that other, non-BAT pathways of thermogenesis are important.

The mitochondrion is the cellular furnace where fuels (dervied from fatty acids and glucose) are oxidized and energy is either stored in the high-energy phosphate bonds of ATP or is released as heat. Figure 3 illustrates that as electrons are passed down, the energy gradient of the electron transport chain, protons are pumped out of the inner matrix of the mitochondria, generating an electrochemical gradient across the inner mitochondrial membrane. These protons have two probable fates: as described by Mitchell, they can reenter the mitochondrial matrix through ATP synthase, driving the synthesis of ATP. ATP production is thus linked to the consumption of oxygen and this is referred to as coupled respiration. Alternatively, protons may "leak" back across the

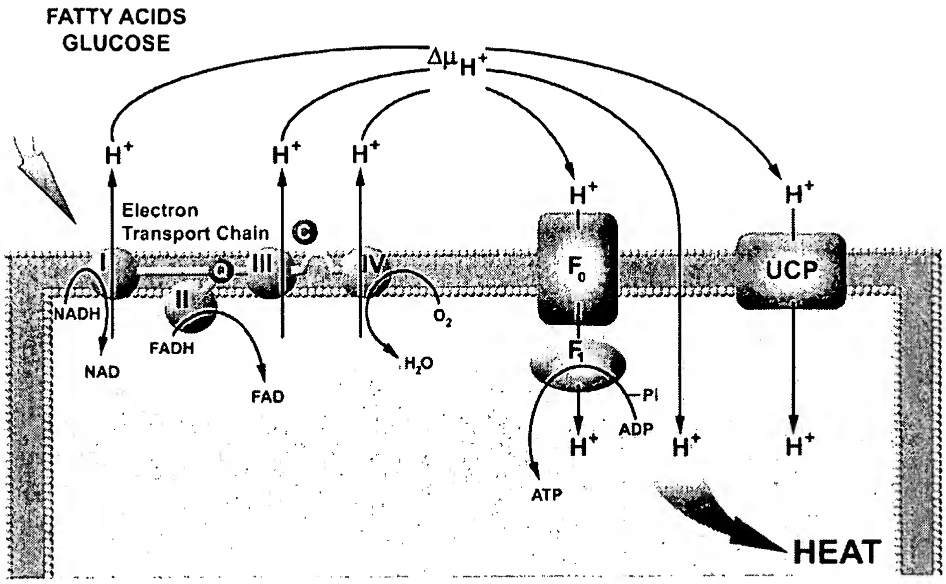

Figure 3. Mitochondrial ATP Metabolism and Thermogenesis through Proton Transport

Fatty acids and glucose are oxidized to generate NADH and $FADH_2$, which donate electrons to the electron transport chain. Ubiquinone (Q) shuttles electrons from both complexes I and II to complex III, whereas cytochrome c (C) shuttles electrons from complex III to complex IV. Molecular oxygen (O_2) is the terminal electron acceptor. Protons are pumped out by complexes I, III, and IV of the electron transport chain, which creates a proton electrochemical potential gradient ($\Delta\mu_{H+}$). Protons may reenter the mitochondrial matrix through the F_0/F_1-ATPase, with energy being used to generate ATP from ADP and Pi. Protons may also reenter through an uncoupling protein (UCP), or the membrane itself, with energy being released in the form of heat. Abbreviations: complex I, NADH-ubiquinone oxidoreductase; complex II, succinate-ubiquinone oxidoreductase; complex III, ubiquinone-cytochrome-c oxidoreductase; and complex IV, cytochrome-c oxidase.

inner mitochondrial membrane in a manner not linked to ATP production. This uncouples energy storage from oxygen consumption and is referred to as uncoupled respiration. A certain degree of "leak" is an inherent property of several biological membranes, but this proton translocation can be greatly accelerated by the action of uncoupling proteins (UCPs), which function as specialized proton channels not linked to ATP production. In uncoupled respiration, energy is released as heat because these leaks, whether catalyzed by UCPs or not, disrupt the cycle and result in fuel oxidation in the absence of work.

The thermogenic function of BAT and UCP-1 have been extensively studied. Interestingly, mice genetically reduced in BAT are prone to obesity while mice deficient in UCP-1 have not shown a propensity to gain weight (Lowell and Flier, 1997). On the other hand, mice lacking UCP-1 are extremely sensitive to cold. These data suggested that there may well be additional mechanisms that can control energy expenditure and metabolic rates in addition to UCP-1, even in BAT.

The last several years have seen the identification of two other members of the UCP family: UCP2, widely expressed in many tissues, and UCP3, expressed primarily in BAT and skeletal muscle. Hopes have been high that these proteins may play major roles in whole body energy expenditure outside of BAT, particularly in skeletal muscle (Kozak and Harper, 2000; Ricquier and Bouillaud, 2000). While it seems clear that these proteins can uncouple oxidative phosphorylation at a cellular level and UCP-3 can protect against obesity when greatly overexpressed in the skeletal muscle of mice (Clapham et al., 2000), several pieces of data have suggested that the role of these proteins in energy balance and physiology may be complex. Most strikingly, fasting, which is associated with a major *reduction* in total body energy expenditure, is associated with an *increase* in the expression of UCP2 and UCP3 mRNA (Boss et al., 2000). More recently, mice deficient in UCP2 or UCP3 have been made and neither strain shows a significant propensity for hypothermia, reduced energy expenditure, or obesity (Arsenijevic et al., 2000; Gong et al., 2000; Vidal-Puig et al., 2000). Together, these data suggest that none of the known UCPs *alone* has clear-cut anti-obesity effects. Of course, it is still entirely possible that they can functionally compensate for each other in targeted mutations, so combined mutations must be created before absolute conclusions can be drawn.

Where does this leave us with regard to mechanisms of energy expenditure and adaptive thermogenesis in large animals like humans, where the mass of skeletal muscle vastly exceeds that of brown fat? There are several possibilities. First, since the inner membrane of mitochondria itself leaks protons, it is possible that adaptive thermogenesis is controlled in an important way by regulation of mitochondrial biogenesis and electron transport rates themselves. An increased electron transport system will drive a greater membrane potential as more protons are pumped out. It is well recognized that this greater membrane potential alone will result in increased proton leak and increase uncoupled respiration even in the absence of a UCP. Hence, UCP1 may be a very specialized molecule for defense against cold in small mammals, but large animals with smaller surface

to volume ratios may simply be able to use the inherent properties of biological membranes to regulate heat production and adaptive thermogenesis.

Alternatively, one can imagine that other "futile-cycles" can serve as specific thermogenic mechanisms regulating metabolic rates (reviewed in Lowell and Spiegelman, 2000). From a bioenergetic perspective, there is nothing magical about a futile cycle of proton pumping that could not, in principle, be accomplished through other means. Key examples exist—deep diving fish have a specially modified periocular muscle that is relatively devoid of cotractile elements. Depolarization of these muscle cells causes release of calcium from the sarcoplasmic reticulum. ATP is then consumed by the Ca^{2+}-ATPase, which pumps calcium back into the sarcoplasmic reticulum in a futile ion cycle. Ca^{2+} cycling results in fuel oxidation without work being performed, thus generating heat. Mammals apparently have a potential to cycle calcium in a similar way. Humans or pigs carrying mutations in the ryanidine receptor, the Ca^{2+} release channel of the sarcoplasmic reticulum, release Ca^{2+} in response to stress or an anesthesia. This results in futile calcium cycling and an intense thermogenesis that can be fatal. Such a mechanism is also observed in cold-adapted birds, suggesting that this could play a role in adaptive thermogenesis of other animals including humans.

Transcriptional Control of Thermogenesis

Several key steps of mammalian thermogenesis, particularly mitochondrial biogenesis and the expression of UCP1, have been the subject of considerable study at the transcriptional level. Much of this effort has been focused on the role of cAMP in inducing these processes.

The UCP1 gene has an enhancer element that is both brown fat selective and responsive to cyclic AMP stimulation (Cassard-Doulcier et al., 1993; Kozak et al., 1994). Perhaps surprisingly, there have been no reports to date of brown fat–specific transcription factors regulating this enhancer. Rather, there are binding sites for PPARγ, an important regulator of both white and brown fat cell differentiation, a potential thyroid hormone response element (TRE), a retinoic acid response element (RARE), and several potential cyclic-AMP response elements (Sears et al., 1996). As defined in knockout mice, the only transcriptional component shown to be required for BAT development is PPARγ (Barak et al., 1999).

The genes of mitochondrial biogenesis and the respiratory chain have been intensively studied, leading to the discovery of nuclear respiratory factor (NRF)-1 and -2 as key transacting elements. The vast majority of mitochondrial genes that are encoded in the nuclear genome have functional binding sites for NRF-1, NRF-2, or both (Virbasius and Scarpulla, 1994; Gugneja et al., 1996). The NRFs, in turn, regulate mitochondrial transcription factor A (mtTFA), which directs the transcription and replication of the mitochondrial genome. How external stimuli, such as cold or diet, affect the amount or activity of NRFs has not been well studied.

Recently, a transcriptional component, PPARγ coactivator (PGC)-1, that can coactivate and coordinate many transcription factors that participate in multiple aspects of adaptive thermogenesis has been described (Puigsever et al., 1998). PGC-1 is expressed in multiple tissues of rodents and man, but is cold inducible only in BAT and skeletal muscle. This cold induction is the result of the sympathetic nervous system acting via β-3 adrenergic receptors and cyclic cAMP (Boss et al., 1999). When PGC-1 is expressed in white fat or skeletal muscle cells, a broad program of thermogenesis begins, including induction of mitochondrial biogenesis, expression of an uncoupling protein (UCP-1 or UCP-2 in fat cells or muscle, respectively), and an increase in total cellular respiration. From the perspective of the adipose lineage, PGC-1 expression makes white fat cells more like brown cells (Puigserver et al., 1998; Wu et al., 1999).

PGC-1 can interact with and coactivate a large number of transcription factors in addition to PPARγ. Indeed, PGC-1 can interact with most nuclear hormone receptors (Puigserver et al., 1998; Knutti et al., 2000; Tsherepanova et al., 2000). PGC-1 docks on some receptors, such as PPARγ or PPARα, in a non-ligand-dependent way and interacts with other receptors, such as the estrogen receptor or glucocorticoid receptor, in a ligand-dependent manner. The ability of PGC-1 to promote mitochondrial biogenesis appears to be due to its ability to turn on the expression of both NRF-1 and NRF-2, and to directly coactivate NRF-1 through protein–protein interactions (Wu et al., 1999). PGC-1 loses most or all of its ability to activate mitochondrial biogenesis in the presence of a dominant-negative allele of NRF-1. The ability of PGC-1 to activate the UCP-1 enhancer appears linked to the coactivation of PPARγ, since PGC-1's activity on this enhancer is ablated by a mutation of the PPARγ binding site.

It is interesting to note that while most biological programs studied to date show dominant regulation at the level of the DNA binding transcription factor, adaptive thermogenesis shows remarkable regulation at the level of the coactivator. Presumably, this is due to the fact that this complex program, which involves multiple tissues, hormone sensitivities, and transcription factors, requires more genetic coordination than can be achieved at the level of a single DNA binding factor. A crucial question that must be answered is whether expression of PGC-1 itself determines or helps to determine whether cells are white fat cells or brown fat cells. Similarly, skeletal muscle fibers come in two types: slow twitch (type 1) that are very oxidative and are very rich in mitochondria, and fast twitch (type 2), which have a more glycolytic metabolism and have fewer mitochondria. Since slow twitch fibers are more oxidative, a greater number of these versus fast twitch fibers could well alter total body energy expenditure. The role of PGC-1 in both of these cell biological decisions must be determined via gain- and loss-of-function experiments in mice.

Linking Lipogenesis and Adipogenesis—A Role for ADD1/SREBP1?

A chronic imbalance between energy intake and energy expenditure can lead to an increase in both fat cell size and fat cell number. Recent studies have suggested a mechanism that can potentially link lipogenesis and adipogenesis.

Several recent reviews have been written on the transcriptional control of adipogenesis (Mandrup and Lane 1997; Rosen et al., 2000), so the details of this process will not be discussed here. What now seems very clear

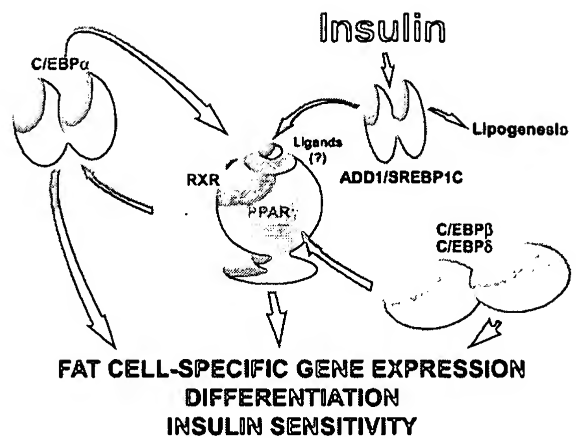

Figure 4. The Transcriptional Control of Adipogenesis Involves the Activation of Several Families of Transcription Factors

These proteins are expressed in a network in which C/EBPβ and C/EBPδ are detected first, followed by PPARγ, which in turn activates C/EBPα and a broad program of adipogenesis. C/EBPα exerts positive feedback on PPARγ to maintain the differentiated state. ADD1/SREBP1c is regulated by insulin in fat and can activate PPARγ by inducing its expression as well as by promoting the production of an endogenous PPARγ ligand. ADD1/SREBP1c also activates many genes of lipogenesis. All of these factors contribute to the expression of genes that characterize the terminally differentiated phenotype.

is that fat cell differentiation is regulated by the function and interplay between several members of the C/EBP transcription factor family, and the nuclear receptor PPARγ. This is summarized in Figure 4. A key question has been, how can a chronic imbalance between energy intake and energy expenditure trigger both fat cell hypertrophy and fat cell hyperplasia? There is no doubt concerning the dominant role played by insulin in the fed and chronically overfed states. There have been a number of transcription factors that can potentially regulate different genes of lipogenesis, but ADD1/SREBP1, a member of the basic helix-loop-helix family of factors, can activate a broad program of genes involved in fatty acid and triglyceride metabolism in both fat and liver (Kim and Spiegelman, 1996; Shimano et al., 1996). Importantly, the expression of ADD1/SREBP1 is regulated by fasting and feeding and this was shown to be regulated by insulin in fat (Kim et al., 1998a). Subsequent studies showed that insulin also regulated this factor in liver (Shimomura et al., 1999). ADD1/SREBP1 has been shown to be synthesized as a membrane-bound precursor that must be released by proteolysis (Brown and Goldstein, 1998; Brown et al., 2000). Cleavage of the ADD1/SREBP1 does not seem to be regulated by cholesterol levels, unlike SREBP2, and the pathway responsible for this activation for ADD1/SREBP1 is still unclear.

ADD1/SREBP1 can also accelerate adipogenesis. This factor alone cannot promote differentiation of non-adipogenic fibroblasts, but when coexpressed on fibroblasts expressing PPARγ, cell differentiation is enhanced (Kim and Spiegelman, 1996). Subsequent studies have shown that ADD1/SREBP1 can enhance the transcriptional activity of PPARγ, and indeed, can activate an isolated ligand binding domain of PPARγ fused to the DNA binding domain of the yeast Gal4 (Kim et al., 1998b). It is most likely that this occurs via transcriptional control of the enzymes required to make an endogenous, agonist ligand of PPARγ. This possibility is further strengthened by the fact that ADD1/SREBP1 controls several known enzymes of fatty acid metabolism, and all known biological ligands for PPARγ are fatty acids or fatty acid derivatives. The identity of the precise PPARγ ligands regulated by ADD1/SREBP1 are unknown but of great interest.

Therapeutic Issues and Opportunities

It is estimated that 300,000 people die annually in the United States as a result of obesity, and most of these deaths are due to the effect of obesity in promoting diabetes, hypertension, cardiovascular disease, and cancer (Kopelman, 2000). Therapy based on nutritional and behavioral counseling is capable of producing useful weight loss with attendant reduction of morbid consequences, but weight loss is usually partial, and almost always temporary. Existing pharmaceuticals target central serotonergic and adrenergic pathways, or inhibit intestinal fat absorption, and are of limited efficacy (Bray and Tartaglia, 2000). The approval of new therapies for obesity will require high standards for safety for several reasons. These include the likely need for chronic therapy, the history of toxicity of serotonergic drugs, concern about abuse by those seeking weight loss for purely cosmetic reasons, and the common, if erroneous view that obese individuals should be able to lose weight through personal effort rather than drugs. Despite these concerns, new insights into the molecular and physiologic pathways that underlie regulated energy balance have created many opportunities for drug discovery in the obesity field.

Obesity Therapy and Energy Balance

A successful obesity therapy must impact energy intake, energy expenditure, or both. Since several key targets (e.g., leptin receptors and MC4Rs) coordinately increase energy expenditure and suppress energy intake, it is possible to envisage therapies that accomplish both effects. Therapies that target energy intake or expenditure alone may initially produce weight loss, but the existence of a homeostatically defended feedback loop would be expected to resist further weight loss and limit efficacy. For example, if weight loss resulted in a fall in leptin levels, this might resist further weight loss through effects on central pathways to increase hunger and decrease energy expenditure. It is possible that combination therapies aimed at two or more distinct steps in the pathway of energy balance might be necessary, as is often the case in disorders such as hypertension and type II diabetes. Finally, we may eventually target specific therapies to individual lesions underlying obesity in specific cases. It is already evident that rare patients lacking leptin respond dramatically to this hormone, while patients lacking the receptor will be totally unresponsive, and patients with common obesity have a limited response.

Therapies Aimed at Suppressing Food Intake

Despite disappointing results of initial clinical trials, it may be possible to find strategies for administering leptin to patients with common obesity that will favor weight

loss or maintain weight loss brought about by other means (Reviewed in Mantzoros and Flier, 2000). The phenomenon of leptin resistance motivates the search for small molecule leptin receptor agonists that might bypass the blood brain barrier, or receptor sensitizers. Several cytokines suppress appetite and promote weight loss. Ciliary neurotrophic factor (CNTF) is a neurocytokine that has been found to reverse obesity in leptin resistant db/db mice, possibly by activating JAK/STAT signaling pathways in leptin responsive neurons (Gloaguen et al., 1997). Therapeutic trials in obesity are underway. Several gut-derived peptides that influence satiety, such as cholecystokinin (CCK) and glucagon-like peptide-1 (GLP-1) (Flint et al., 1998), or small molecule agonists for their receptors, are therapeutic candidates. Perhaps the greatest pharmaceutical effort is addressing the central neuropeptide pathways. Knockout experiments for NPY and its several receptors suggest a robust capacity for redundancy of these pathways, but small molecule antagonists for NPY Y-1 and Y-5 receptors have shown preclinical promise, and are under development (Gehlert, 1999). It should be stressed that the phenotypes of gene knockout mice may not always predict the response to inhibitors of the same pathway. Although less advanced than NPY antagonists, great effort is being expended to develop MC4R agonists and antagonists for the MCH receptor.

Therapies Aimed at Increasing Energy Expenditure

The fall in energy expenditure during weight loss limits the efficacy of diets, and therapies that could prevent this, or simply increase energy expenditure, would promote weight loss. Thyroid hormone promotes thermogenesis by as yet uncertain mechanisms, and it is well known to clinicians that increased thyroid hormone produces weight loss, but other adverse effects of thyroid excess including loss of lean body mass prevent its use as an obesity treatment. β-3 adrenergic agonists are highly effective at promoting thermogenesis and weight loss in animals, and if selectivity for β-3 receptors can be established, such drugs may be very effective (Himms-Hagen et al., 1994). Likewise, drugs capable of activating or increasing expression of the newly identified mitochondrial uncoupling proteins UCP-2 and UCP-3, or the transcriptional regulator PGC-1, would be interesting therapeutic candidates.

Therapies that May Limit Obesity by Uncertain Mechanisms

The results of several gene knockout experiments have produced mice with resistance to obesity that was not anticipated. Protein tyrosine phosphatase-1B (PTP 1-B) knockout mice have enhanced insulin sensitivity, and are also resistant to obesity caused by high-fat diet (Klaman et al., 2000). Although the mechanism for the effect on energy balance is uncertain, PTP-1B could be a target for a drug to treat both diabetes and obesity. The enzyme acyl CoA:diacylglycerol transferase (DGAT) mediates the final step in the glycerol phosphate pathway of triglyceride synthesis. Mice lacking DGAT are lean and resistant to obesity induced by high-fat diet, and surprisingly manifest increased energy expenditure that remains unexplained (Smith et al., 2000). DGAT may also be a target for anti-obesity therapy. Perilipin is an adipocyte protein that regulates lipolysis through effects on hormone sensitive lipase. A perilipin knockout mouse, in which adipocyte lipase is constitutively increased, is lean despite increased food intake, and has increased metabolic rate; absence of perilipin dramatically reduces obesity of db/db (Martinez-Botas et al., 2000). Although some aspects of the phenotype, such as increased energy expenditure, are yet to be explained, perilipin may be an interesting drug target.

Conclusions

Obesity and its antithesis, starvation, have always been part of the human condition, and for most of human history have been seen as resulting simply from availability of food, or acts of will related to attainment of desired body shape. Although this view persists in some quarters to this day, the last 5 years of the millennium have witnessed a dramatic increase in our understanding of the biology of regulated energy balance and body weight. Physiologic pathways whose existence was debated 10 years ago are now being characterized in molecular detail, with immediate implications for understanding of pathogenesis of human obesity and other disorders of energy balance. The roadmap provided by these advances establishes a clear direction for future research, but critical details remain to be discovered, and therapeutic applications remain to be realized. In particular, the mechanisms by which environmental factors, including diet and exercise, interact with molecular pathways in the common polygenic forms of obesity is largely unknown at present. Insights from the sequencing of the human genome and the coming advances in proteomics are likely to fuel the next wave of progress. It is likely that both new genes and new regulatory pathways will be identified. It may seem unlikely that the recent wave of progress can be matched in the early years of the current millennium, but we would not choose to make that bet.

References

Ahima, R.S., and Flier, J.S. (2000). Leptin. Annu. Rev. Physiol. *62*, 413–437.

Ahima, R.S., Prabakara, D., Mantzoros, C., Qu, D., Lowell, B., Maratos-Flier, E., and Flier, J.S. (1996). Role of leptin in the neuroendocrine response to fasting. Nature *82*, 250–252.

Arsenijevic, D., Onuma, H., Pecqueur, C., Raimbault, S., Manning, B.S., Miroux, B., Couplan, E., Alves-Guerra, M.C., Goubern, M., Surwit, R., et al. (2000). Disruption of the uncoupling protein-2 gene in mice reveals a role in immunity and reactive oxygen species production. Nat. Genet. *26*, 435–439.

Barak, Y., Nelson, M.C., Ong, E.S., Jones, Y.Z., Ruiz-Lozano, P., Chien, K.R., Koder, A., and Evans, R.M. (1999). PPAR gamma is required for placental, cardiac, and adipose tissue development. Mol. Cell *4*, 585–595.

Barsh, G.S., Farooqi, I.S., and O'Rahilly, S. (2000). Genetics of body-weight regulation. Nature *404*, 644–651.

Billington, C.J., Briggs, J.E., Harker, S., Grace, M., and Levine, A.S. (1994). Neuropeptide Y in hypothalamic paraventricular nucleus: a center coordinating energy metabolism. Am. J. Physiol. *266*, R1765–R1770.

Bjorbaek, C., Elmquist, J.K., Frantz, J.D., Shoelson, S.E., and Flier, J.S. (1998). Identification of SOCS-3 as a potential mediator of central leptin resistance. Mol. Cell *191*, 619–625.

Bjorbaek, C., El-Haschimi, K., Frantz, J.D., and Flier, J.S. (1999). The role of SOCS-3 in leptin signaling and leptin resistance. J. Biol. Chem. *274*, 30059–30065.

Bjorbaek, C., Lavery, H.J., Bates, S.H., Olson, R.K., Davis, S.M., Flier, J.S., and Myer, M.G., Jr. (2000). SOCS3 mediates feedback inhibition of the leptin receptor via Tyr985. J. Biol. Chem., in press.

Boss, O., Bachman, E., Vidal-Puig, A., Zhang, C.Y., Peroni, O., and Lowell, B.B. (1999). Role of the beta(3)-adrenergic receptor and/or a putative beta(4)-adrenergic receptor on the expression of uncoupling proteins and peroxisome proliferator-activated receptor-gamma coactivtor-1. Biochem. Biophys. Res. Commun. *261*, 870–876.

Boss, O., Hagen, T., and Lowell, B.B. (2000). Uncoupling proteins 2 and 3: potential regulators of mitochondrial eneergy metabolism. Diabetes *49*, 143–156.

Bouchard, C., Tremblay, A., Despres, J.P., Nadeau, A., Lupien, P.J., Theriault, G., Dussault, J., Moorjani, S., Pinault, S., and Fournier, G. (1990). The response to long-term overfeeding in identical twins. N. Engl. J. Med. *322*, 1477–1482.

Bray, G.A., and Tartaglia, L.A. (2000). Medicinal strategies in the treatment of obesity. Nature *404*, 672–677.

Brown, M.S., and Goldstein, J.L. (1998). Sterol regulatory element binding proteins (SREBPs): controllers of lipid synthesis and cellular uptake. Nutr. Rev. *56*, S54–S75.

Brown, M.S., Ye, J., Rawson, R.B., and Goldstein, J.L. (2000). Regulated intramembrane proteolysis: a control mechanism conserved from bacteria to humans. Cell *100*, 391–398.

Bruning, J.C., Gautam, D., Burks, D.J., Gillette, J., Schubert, M., Orban, P.C., Klein, R., Krone, W., Muller-Wieland, D., and Kahn, C.R. (2000). Role of brain insulin receptor in control of body weight and reproduction. Science *289*, 2122–2125.

Bultman, S.J., Michaud, E.J., and Woychik, R.P. (1992). Molecular characterization of the mouse agouti locus. Cell *71*, 1195–1204.

Burks, D.J., de Mora, J.F., Schubert, M., Withers, D.J., Myers, M.G., Towery, H.H., Altamuro, S.L., Flint, C.L., and White, M.F. (2000). IRS-2 pathways integrate female reproduction and energy homeostasis. Nature *407*, 377–382.

Butler, A.A., Kesterson, R.A., Khong, K., Cullen, M.J., Pelleymounter, M.A., Dekoning, J., Baetscher, M., and Cone, R.D. (2000). A unique metabolic syndrome causes obesity in the melanocortin-3-receptor-deficient mouse. Endocrinology *141*, 3518–3521.

Cassard-Doulcier, A.M., Gelly, C., Fox, N., Schrementi, J., Raimbault, S., Klaus, S., Forest, C., Bouillard, F., and Ricquier, D. (1993). Tissue-specific and beta-adrenergic regulation of the mitochondrial uncoupling protein gene: control by cis-acting elements in the 5′-flanking region. Mol. Endocrinol. *7*, 497–506.

Chen, H., Charlat, O., Tartaglia, L.A., Woolf, E.A., Weng, X., Ellis, S.J., Lakey, N.D., Culpepper, J., Moore, K.J., Breitbart, R.E., et al. (1996). Evidence that the diabetes gene encodes the leptin receptor: identification of a mutation in the leptin receptor gene in db/db mice. Cell *84*, 491–495.

Clapham, J.C., Arch, J.R., Chapman, H., Haynes, A., Lister, C., Moore, G.B., Piercy, V., Carter, S.A., Lehner, I., Smith, S.A., et al. (2000). Mice overexpressing human uncoupling protein-3 in skeletal muscle are hyperphagic and lean. Nature *406*, 415–418.

Comuzzie, A.G., Hixson, J.E., Almasy, L., Mitchell, B.D., Mahaney, M.C., Dyer, T.D., Stern, M.P., MacCluer, J.W., and Blangero, J. (1997). A major quantitative trait locus determining serum leptin levels and fat mass is located on human chromosome 2. Nat. Genet. *15*, 273–276.

Cone, R.D. (1999). The central melanocortin system and energy homeostasis. Trends Endocrinol. *10*, 211–216.

Considine, R.V., Sinha, M.K., Heiman, M.L., Kriauciunas, A., Stephens, T.W., Nyce, M.R., Ohannesian, J.P., Marco, C.C., McKee, L.J., and Bauer, T.L. (1996). Serum immunoreactive-leptin concentrations in normal-weight and obese humans. N. Engl. J. Med. *334*, 292–295.

Cowley, M.A., Pronchuk, N., Fan, W., Dinulescu, D.M., Colmers, W.F., and Cone, R.D. (1999). Integration of NPY, AGRP, and melanocortin signals in the hypothalamic paraventricular nucleus: evidence of a cellular basis for the adipostat. Neuron *24*, 155–163.

Ducy, P., Amling, M., Takeda, S., Priemel, M., Schilling, A.F., Beil, F.T., Shen, J., Vinson, C., Ruegger, J.M., and Karsenty, G. (2000). Leptin inhibits bone formation through a hyppothalamic relay: a central control of bone mass. Cell *100*, 197–207.

El-Haschimi, K., Pierroz, D.D., Hileman, S.M., Bjorbaek, C., and Flier, J.S. (2000). Two defects contribute to hypothalamic leptin resistance in mice with diet-induced obesity. J. Clin. Invest. *105*, 1827–1832.

Elias, C.F., Lee, C., Kelly, J., Aschkenasi, C., Ahima, R.S., Couceyro, P.R., Kuha, M.J., Saper, C.B., and Elmquist, J.K. (1998). Leptin activates hypothalamic CART neurons projecting to the spinal cord. Neuron *21*, 1375–1385.

Elias, C.F., Aschkenasi, C., Lee, C., Kelly, J., Ahima, R.S., Bjorbaek, C., Flier, J.S., Saper, C.B., and Elmquist, J.K. (1999). Leptin differentially regulates NPY and POMC neurons projecting to the lateral hypothalamic area. Neuron *23*, 775–786.

Elmquist, J.K., Elias, C.F., and Saper, C.B. (1999). From lesions to leptin: hypothalamic control of food intake and body weight. Neuron *22*, 221–232.

Emanuelli, B., Peraldi, P., Filloux, C., Sawka-Verhelle, D., Hilton, D., and Van Obberghen, E. (2000). SOCS-3 is an insulin-induced negative regulator of insulin signaling. J. Biol. Chem. *275*, 15985–15991.

Emond, M., Schwartz, G.J., Ladenheim, E.E., and Moran, T.H. (1999). Central leptin modulates behavioral and neural responsivity to CCK. Am. J. Physiol. *276*, R1545–R1549.

Enerback, S., Jacobsson, A., Simpson, E.M., Guerra, C., Yamashita, H., Harper, M.E., and Kozak, L.P. (1997). Mice lacking mitochondrial uncoupling protein are cold-sensitive but not obese. Nature *387*, 90–94.

Erickson, J.C., Clegg, K.E., and Palmiter, R.D. (1996a). Sensitivity to leptin and susceptibility to seizures of mice lacking neuropeptide Y. Nature *381*, 415–421.

Erickson, J.C., Hollopeter, G., and Palmiter, R.D. (1996b). Attenuation of the obesity syndrome of ob/ob mice by the loss of neuropeptide Y. Science *274*, 1704–1707.

Fan, W., Boston, B.A., Kesterson, R.A., Hruby, V.J., and Cone, R.D. (1997). Role of melanocortinergic neurons in feeding and the agouti obesity syndrome. Nature *385*, 165–168.

Farooqi, I.S., Yeo, G.S., Keogh, J.M., Aminian, S., Jebb, S.A., Butler, G., Cheetham, T., and O'Rahilly, S. (2000). Dominant and recessive inheritance of morbid obesity associated with melanocortin 4 receptor deficiency. J. Clin. Invest. *106*, 271–279.

Flier, J.S., Harris, M., and Hollenberg, A.N. (2000). Leptin, nutrition, and the thyroid: the why, the wherefore, and the wiring. J. Clin. Invest. *105*, 859–861.

Flint, A., Raben, A., Astrup, A., and Holst, J.J. (1998). Glucagon-like peptide 1 promotes satiety and suppresses energy intake in humans. J. Clin. Invest. *101*, 515–520.

Friedman, J.M., and Halaas, J.L. (1998). Leptin and the regulation of body weight in mammals. Nature *395*, 763–770.

Gehlert, D.R. (1999). Role of hypothalamic neuropeptide Y in feeding and obesity. Neuropeptides *33*, 329–338.

Gloaguen, I., Costa, P., Demartis, A., Lazzaro, D., DiMarco, A., Graziani, R., Paonessa, G., Chen, F., Rosenblum, C.I., Van der Ploeg, L.H., et al. (1997). Ciliary neurotrophic factor corrects obesity and obesity associated with leptin deficiency and resistance. Proc. Natl. Acad. Sci. USA *94*, 6456–6461.

Gong, D.W., Monemdjou, S., Gavrilova, O., Leon, L.R., Marcus-Samuels, B., Chou, C.J., Everett, C., Kozak, L.P., Li, C., Deng, C., et al. (2000). Lack of obesity and normal response to fasting and thyroid hormone in mice lacking uncoupling protein-3. J. Biol.Chem. *275*, 16251–16257.

Graham, M., Shutter, J.R., Sarmiento, U., Sarosi, I., and Stark, K.L. (1997). Overexpression of Agrp leads to obesity in transgenic mice. Nat. Genet. *17*, 273–274.

Gugneja, S., Virbasius, C.M., and Scarpulla, R.C. (1996). Nuclear respiratory factors 1 and 2 utilize similar glutamine-containing clusters of hydrophobic residues to activate transcription. Mol. Cell. Biol. *16*, 5708–5716.

Halaas, J.L., Gajiwala, K.S., Maffei, M., Cohen, S.L., Chait, B.T., Rabinowitz, D., Lallone, R.L., Burley, S.K., and Friedman, J.M. (1995).

Weight-reducing effects of the plasma protein encoded by the obese gene. Science *269*, 543–546.

Himms-Hagen, J., Cui, J., Danforth, E., Jr., Taatjes, D.J., Lang, S.S., Waters, B.L., and Claus, T.H. (1994). Effect of CL-316,243, a thermogenic beta 3-agonist, on energy balance and brown and white adipose tissues in rats. Am. J. Physiol. *266*, R1371–R1382.

Huszar, D., Lynch, C.A., Fairchild-Huntress, V., Dunmore, J.H., Fang, Q., Berkemeier, L.R., Gu, W., Kesterson, R.A., Boston, B.A., Cone, R.D., et al. (1997). Targeted disruption of the melanocortin-4 receptor results in obesity in mice. Cell *88*, 131–141.

Kim, J.B., and Spiegelman, B.M. (1996). ADD1/SREBP1 promotes adipocyte differentiation and gene expression linked to fatty acid metabolism. Genes Dev. *10*, 1096–1107.

Kim, J.B., Sarraf, P., Wright, M., Mueller, E., Lowell, B.B., and Spiegelman, B.M. (1998a). Nutritional and insulin regulation of fatty acid synthetase and leptin gene expression through ADD1/SREBP1. J. Clin. Invest. *101*, 1–9.

Kim, J.B., Wright, H.M., Wright, M., and Spiegelman, B.M. (1998b). ADD1/SREBP1 activates PPARγ through the production of endogenous ligand. Proc. Natl. Acad. Sci. USA *95*, 4333–4337.

Klaman, L.D., Boss, O., Peroni, O.D., Kim, J.K., Martino, J.L., Zabolotny, J.M., Moghal, N., Lubkin, M., Kim, Y.B., Sharpe, A.H., et al. (2000). Increased energy expenditure, decreased adiposity, and tissue-specific insulin sensitivity in protein-tyrosine phosphatase 1B-deficient mice. Mol. Cell. Biol. *20*, 5479–5489.

Knutti, D., Kaul, A., and Kralli, A. (2000). A tissue-specific coactivator of steroid receptors, identified in a functional genetic screen. Mol. Cell. Biol. *20*, 2411–2422.

Kopelman, P.G. (2000). Obesity as a medical problem. Nature *404*, 635–643.

Kozak, U.C., Kopecky, J., Teisinger, J., Eneback, S., Boyer, B., and Kozak, L.P. (1994). An upstream enhancer regulating brown fat-specific expression of the mitochondrial uncoupling protein gene. Mol. Cell. Biol. *14*, 59–67.

Kozak, L.P., and Harper, M.E. (2000). Mitochondrial uncoupling proteins in energy expenditure. Ann .Rev. Nature *20*, 339–363.

Krude, H., Biebermann, H., Luck, W., Horn, R., Brabant, G., and Gruters, A. (1998). Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. Nat. Genet. *19*, 155–157.

Kuczmarski, R.J., Flegal, K.M., Campbell, S.M., and Johnson, C.L. (1994). Increasing prevalence of overweight among US adults. The National Health and Nutrition Examination Surveys, 1960 to 1991. JAMA *272*, 205–211.

Lee, Y., Wang, M.Y., Kakuma, T., Wang, Z.W., Babcock, E., McCorkle, K., Higa, M., Zhou, Y.T., and Unger, R.H. (2000). Liporegulation in diet-induced obesity: the antisteatotic role of hyperleptinemia. J. Biol. Chem., in press. 10.1074/jbc.M008553200.

Lichtman, S.W., Pisarska, K., Berman, E.R., Pestone, M., Dowling, H., Offenbacher, E., Weisel, H., Heshka, S., Matthews, D.E., and Heymsfield, S.B. (1992). Discrepancy between self-reported and actual caloric intake an excerise in obese subjects. N. Eng. J. Med. *327*, 1893–1898.

Loftus, T.M., Jaworsky, D.E., Frehywot, G.L., Townsend, C.A., Ronnett, G.V., Lane, M.K., and Kuhajda, F.P. (2000). Reduced food intake and body weight in mice treated with fatty acid synthase inhibitors. Science *288*, 2379–2381.

Lowell, B.B., and Flier, J.S. (1997). Brown adipose tissue, beta 3-adrenergic receptors, and obesity. Ann. Rev. Med. *48*, 307–316.

Lowell, B.B., and Spiegelman, B.M. (2000). Towards a molecular understanding of adaptive thermogenesis. Nature *404*, 652–660.

Lu, D., Willard, D., Patel, I.R., Kandell, S., Overton, L., Kost, T., Luther, M., Chen, W., Woychick, R.P., Wilkison, W.O., and Cone, R. (1994). Agouti protein is an antagonist of the melanocyte stimulating hormone receptor. Nature *371*, 799–802.

Mandrup, S., and Lane, M.D. (1997). Regulating adipogenesis. J. Biol. Chem. *272*, 5367–5370.

Mantzoros, C.S., and Flier, J.S. (2000). Editorial: leptin as a therapeutic agent-trials and tribulations. J. Clin. Endocrinol. Metab. *85*, 4000–4002.

Martinez-Botas, J., Anderson, J.B., Tessier, D., Lapillonne, A., Chang, B.H., Quast, M.J., Gorenstein, D., Chen, K.H., and Chan, L. (2000). Absence of perilipin results in leanness and reverses obesity in Leprdb/db mice. Nat. Genet. *26*, 474–479.

Moitra, J., Mason, M.M., Olive, M., Krylov, D., Gavrilova, O., Marcus-Samuels, B., Feigenbaum, L., Lee, E., Aoyama, T., Eckhaus, M., et al. (1998). Genes Dev. *12*, 168–181.

Moran, T.H. (2000). Cholecystokinin and satiety: current perspectives. Nutrition *16*, 858–865.

Neel, J.V. (1999). The "thrifty genotype" in 1998. Nutr. Rev. *57*, S2-9.

Nonogaki, K., Strack, A.M., Dallman, M.F., and Tecott, L.H. (1998). Leptin-independent hyperphagia and type 2 diabetes in mice with a mutated serotonin 5-HT2C receptor gene. Nat. Med. *4*, 1152–1156.

Ollmann, M.M., Wilson, B.D., Yang, Y.J., Kerns, J.A., Chen, Y., Gantz, I., and Barsh, G.S. (1997). Antagonism of central melanocortin receptors in vitro and in vivo by agouti-related protein. Science *278*, 135–138.

Puigserver, P., Wu, Z., Park, C.W., Graves, R., Wright, M., and Spiegelman, B.M. (1998). A cold inducible coactivator of nuclear receptors linked to adaptive thermogenesis. Cell *92*, 829–839.

Qu, D., Ludwig, D.S., Gammeltoft, S., Piper, M., Pelleymounter, M.A., Cullen, M.J., Mathes, W.F., Przypek, R., Kanarek, R., and Maratos-Flier, E. (1996). A role for melanin-concentrating hormone in the central regulation of feeding behaviour. Nature *380*, 243–247.

Ravussin, E. (1995). Metabolic differences and the development of obesity. Metabolism *44*, 12–14.

Ricquier, D., and Bouillaud, F. (2000). Mitochondrial uncoupling proteins: from mitochondria to the regulation of energy balance. J. Physiol. *1*, 3–10.

Rosen, E.D., Walkey, C.J., Puigserver, P., and Spiegelman, B.M. (2000). Transcriptional regulation of adipogenesis. Genes Dev. *14*, 1293–1307.

Ross, S.R., Graves, R.A., and Spiegelman, B.M. (1993). Targeted expression of a toxin gene to adipose tissue: transgenic mice resistant to obesity. Genes Dev. *7*, 1318–1324.

Schwartz, M.W., Woods, S.C., Porte, D., Jr., Seeley, R.J., and Baskin, D.G. (2000). Central nervous system control of food intake. Nature *404*, 661–671.

Sears, I.B., MacGinnitie, M.A., Kovacs, L.G., and Graves, R.A. (1996). Differentiation-dependent expression of the brown adipocyte uncoupling protein gene: regulation by peroxisome proliferator-activated receptor gamma. Mol. Cell. Biol. *16*, 3410–3419.

Shimada, M., Tritos, N.A., Lowell, B.B., Flier, J.S., and Maratos-Flier, E. (1998). Mice lacking melanin-concentrating hormone are hypophagic and lean. Nature *396*, 670–674.

Shimano, H., Horton, J.D., Hammer, R.E., Shimomura, I., Brown, M.S., and Goldstein, H. (1996). Overproduction of cholesterol and fatty acids causes massive liver enlargement in transgenic mice expressing truncated SREBP-1a. J. Clin. Invest. *98*, 1575–1584.

Shimomura, I., Hammer, R.E., Richardson, J.A., Ikemoto, S., Bashmakov, Y., Goldstein, J.L., and Brown, M.S. (1998). Insulin resistance and diabetes mellitus in transgenic mice expressing nuclear SREBP-1c in adipose tissue: model for congenital generalized lipodystrophy. Genes Dev. *12*, 3182–3194.

Shimomura, I., Bashmakov, Y., Ikemoto, S., Horton, J.D., Brown, M.S., and Goldstein, J.L. (1999). Insulin seletively increases SREBP-1c mRNA in the livers of rats with streptozotocin-induced diabetes. Proc. Natl. Acad. Sci. USA *96*, 13656–13661.

Simonsen, I., Bulow, J., Madsen, J., and Christensen, N.J. (1992). Thermogenic response to epinephrine in the forearm and abdominal subcutaneous adipose tissue. Am. J. Physiol. *263*, E850–E855.

Smith, G.P. (2000). The controls of eating: a shift from nutritional homeostasis to behavioral neuroscience. Nutrition *16*, 814–820.

Smith, S.J., Cases, S., Jensen, D.R., Chen, H.C., Sande, E., Tow, B., Sanan, D.A., Raber, J., Eckel, R.H., and Farese, R.V., Jr. (2000). Obesity resistance and multiple mechanisms of triglyceride synthesis in mice lacking Dgat. Nat. Genet. *25*, 87–90.

Stanley, B.G., Kyrkouli, S.E., Lampert, S., and Leibowitz, S.F. (1986). Neuropeptide Y chronically injected into the hypothalamus: a powerful neurochemical inducer of hyperphagia and obesity. Peptides *7*, 1189–1192.

Tartaglia, L.A. (1997). The leptin receptor. J. Biol. Chem. *272*, 6093–6096.

Thomas, S.A., and Palmiter, R.D. (1997). Thermoregulatory and metabolic phenotypes of mice lacking noradrenaline and adrenaline. Nature *387*, 94–97.

Troiano, R.P., and Flegal, K.M. (1999). Overweight prevalence among youth in the United States: why so many different numbers? Int. J. Obes. Relat. Metab. Disord. *2*, S22–S27.

Tschop, M., Smiley, D.L., and Heiman, M.L. (2000). Ghrelin induces adiposity in rodents. Nature *407*, 908–913.

Tsherepanova, I., Puigserver, P., Norris, J.D., Spiegelman, B.M., and McDonnell, D.P. (2000). Modulation of estrogen receptor-alpha transcriptional activity by the coactivator PGC-1. J. Biol. Chem. *275*, 16302–16308.

Van Heek, M., Compton, D.S., France, C.F., Tedesco, R.P., Fawzi, A.B., Graziano, M.P., Sybertz, E.J., Strader, C.D., and Davis, H.R., Jr. (1997). Diet-induced obese mice develop peripheral, but not central, resistance to leptin. J. Clin. Invest. *99*, 385–390.

Vidal-Puig, A.J., Grujic, D., Zhang, C.Y., Hagen, T., Boss, O., Ido, Y., Szczepanik, A., Wade, J., Mootha, V., Cortright, R., et al. (2000). Energy metabolism in uncoupling protein 3 gene knockout mice. J. Biol. Chem. *275*, 16258–16266.

Virbasius, J.V., and Scarpulla, R.C. (1994). Activation of the human mitochondrial transcription factor A gene by nuclear respiratory factors: a potential regulatory link between nuclear and mitochondrial gene expression in organelle biogenesis. Proc. Natl. Acad. Sci. USA *91*, 1309–1313.

Wang, J., Liu, R., Hawkins, M., Barzilai, N., and Rossetti, L. (1998). A nutrient-sensing pathway regulates leptin gene expression in muscle and fat. Nature *393*, 684–688.

Wu, Z., Puigserver, P., Andersson, U., Zhang, C., Adelmant, G., Mootha, V., Troy, A., Cinti, S., Lowell, B., Scarpulla, R., and Spiegelman, B.M. (1999). Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. Cell *98*, 115–124.

Zhang, Y., Proenca, R., Maffei, M., Barone, M., Leopold, L., and Friedman, J.M. (1994). Positional cloning of the mouse obese gene and its human homologue. Nature *372*, 425–432.